

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2005, 18:03:57 ; Search time 4813 Seconds
(without alignments)
11114.598 Million cell updates/sec

Title: US-10-071-645-1
Perfect score: 1104
Sequence: 1 ggcacgagcgccgctg.....aaaaaaaaaaaaaaaaaaaa 1104

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_to.*
11: gb_str.*
12: gb_by.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1104	100.0	1104	6	AX451459 Sequence
2	803.8	72.8	1876	6	AR454597 Sequence
3	790	71.6	1581	6	BD276315 MOLECULES
4	790	71.6	1581	6	AX049462 Sequence
5	632.6	57.3	1321	6	CQ776764 Sequence
6	607.2	55.0	3189	6	AX451461 Sequence
7	607.2	55.0	126141	2	AL356300 Homo sapi
8	607.2	55.0	195076	9	AL391357 Homo sapi
9	605	54.8	1377	9	BC020630 Homo sapi
10	600.6	54.4	163801	2	AC025224 Homo sapi
11	441	39.9	860	9	AY204901 Homo sapi
12	427.2	38.7	524	6	AX894800 Sequence
13	427.2	38.7	524	6	BD030333 Sequence
14	405.4	36.7	417	6	BD139338 Extended
15	343.8	31.1	366	6	CQ704178 Sequence
16	295.8	25.8	360	6	CQ699651 Sequence
17	282.2	25.6	411	6	AX330168 Sequence
18	282.2	25.6	411	6	AX332831 Sequence
19	282.2	25.6	411	6	AX334667 Sequence

20	282.2	25.6	411	6	AX335116 Sequence
21	279.4	25.3	409	10	AF271156 Rattus no
22	273.4	24.8	1592	10	AY523601 Mus muscu
c 23	270	24.5	417	6	AR272746 Sequence
c 24	270	24.5	417	6	AR276327 Sequence
c 25	270	24.5	417	6	AR406602 Sequence
c 26	270	24.5	417	6	AR440452 Sequence
c 27	270	24.5	417	6	AR472610 Sequence
c 28	270	24.5	417	6	AR543263 Sequence
c 29	270	24.5	417	6	AX062848 Sequence
c 30	270	24.5	417	6	AX367765 Sequence
c 31	231.6	21.0	259224	2	AC109006 Rattus no
c 32	226.8	20.5	170878	2	AC107845 Mus muscu
33	226.8	20.5	248479	10	AL807249 Mouse DNA
34	225.6	20.4	892	11	G29976 G29976 human STR S
35	224.4	20.3	237	6	CQ727202 Sequence
c 36	202	18.3	616	9	HS324422 Homo sapi
37	165	14.9	336	6	AR563282 Sequence
38	165	14.9	336	6	AX322029 Sequence
39	155.8	14.1	183	6	AX894673 Sequence
40	155.8	14.1	183	6	BD030206 Sequence
41	133.2	12.1	695	5	CR407503 Gallus ga
42	133.2	11.2	308	6	CQ427555 Sequence
43	133.2	11.2	333	6	CQ418653 Sequence
44	120.4	10.9	347	10	AF041854 Rattus no
45	118.8	10.8	550	6	AX392227 Sequence

ALIGNMENTS

RESULT 1
AX451459
LOCUS AX451459 1104 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent WO0212279.
ACCESSION AX451459
VERSION AX451459.1 GI:21698460
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bihain,B., Bour,B. and Bougueleret,L.
TITLE Schizophrenia related gene and protein
JOURNAL Patent: WO 0212279-A 1 14-FEB-2002;
GENSET (FR)
FEATURES
source
1..1104
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
87..344
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37813.1"
/db_xref="GI:21698461"
/translation="MWVLPYDGDKLSPYGDGVDVGQIFSCRLODTNNFFGAGQNRKP
PKLQIGRSKRVIEDDRIDDLVKNWTDKAPLVNSPKTMS"

ORIGIN
Query Match 100.0%; Score 1104; DB 6; Length 1104;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCACGAGCGAGCGCGCTGACCTGTCCGCGCGCGCGGAGCGCGGAGAGCGG 60
Db 1 GGCACGAGCGAGCGCGCTGACCTGTCCGCGCGCGGAGCGCGGAGAGCGG 60
Qy 61 CCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGC 120
Db 61 CCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGC 120

Qy	121	181	241	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141
Db	121	181	241	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141
Qy	181	241	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201
Db	181	241	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201
Qy	241	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261
Db	241	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261
Qy	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321
Db	301	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321
Qy	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381
Db	361	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381
Qy	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441
Db	421	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441
Qy	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501
Db	481	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501
Qy	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561
Db	541	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561
Qy	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621
Db	601	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621
Qy	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621	1681
Db	661	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621	1681
Qy	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621	1681	1741
Db	721	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621	1681	1741
Qy	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621	1681	1741	1801
Db	781	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621	1681	1741	1801
Qy	841	901	961	1021	1081	1141	1201	1261	1321	1381	1441	1501	1561	1621	1681	1741	1801	1861
Db	84																	

RESULT 2			
AR454597			
LOCUS	AR454597	1876 bp	DNA
DEFINITION	Sequence 70 from patent US 6682888.		linear
			PAT 20-FEB-2004

ACCESSION	AR454597
VERSION	AR454597.1 GI:42688066
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1876)
TITLE	Loring,J.F., Tingley,D.W. and Edwards,C.M.
JOURNAL	Genes expressed in alzheimer's disease
FEATURES	Patent: US 6682868-A 70 27-JAN-2004;
source	Location/Qualifiers
	1..1876
	/organism="unknown"
ORIGIN	/mol_type="genomic DNA"
	Query Match 72.8%; Score 803.8; DB 6; Length 1876;
	Best Local Similarity 87.9%; Pred No. 9.8e-129;
	Matches 991; Conservative 0; Mismatches 103; Indels 34; Gaps 11;
QY	10 CAGCGCCGCTGACCCTGTGCCCGGGCGGGGACGCGGGCGGAGGCGCCCGCGGG 69
Dd	279 CAGCGCCGCTGACCCTGTGCCCGGGCGGGGACGCGGGCGGAGGCGCCCGCGGG 338
QY	70 AGCCCCGGGACCGCACCATGTGGGAGTGTCGCCTACGGCGGACGAGAGCTGAGCCCT 129
Dd	339 AG-CCCCGGGACCGCACCATGTTCGGAGGTGTCGCCTACGGCGGACGAGAAGCTGAGCCCT 397
QY	130 ACGCGCAGCGGGCGGACGTGGGGCCAGATCTTCTCTCGCGCTCGACGACACCAACAAC 189
Dd	398 ACGCGCAGCGGGCGGACGTGGGGCCAGATCTTCTCTCGCGCTCGACGACACCAACAAC 457
QY	190 TC'TTGGGGCGGGCAGAAACGAGCGGGCGCCCAAGCTGGGCCAGATCGGCCGAGCAAGC 249
Dd	458 TC'TTGGGGCGGGCAGAACGAAGCGGGCGCCCAAGCTGGGCCAGATCGGCCGAGCAAGC 517
QY	250 GGGTTGTTATTGAAGATGATAGGATTGATGACGTGCTGMAAATAATGACCGACAAAGGCAC 309
Dd	518 GGGTTGTTATTGAAGATGATAGGATTGATGACGTGCTGMAAANAATGACCGACAAAGGCAC 577
QY	310 CT-CTGGTGCTAACT--CCCCAAAGACAATGAGTTAAGGAGAGAAATAGGAACGGCGGTA 367
Dd	578 CTCCTGGTGCTAACTCTCCCCAAAGACAATGAGTTAAGGAGAGAAATAAGAACGGCGGTA 637
QY	368 ACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACITTTGAAATTTTATTACTAGCTTG-T 426
Dd	638 ACAGTTATTGGCAAAAAGCATGAAAAGAGAAAGCACITTTGNAATTTATTACTAGCTTGCT 697
QY	427 ACCCAGATGAATCAAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGCGGAAG 485
Dd	698 ACCCAGATGAATCAAACAACCTGTATCTGGTATCAGGCCGGGAGACAGATGAGCGGAGA 757
QY	486 GAGCAAGAGAGAGAGAAAGGCTTGGGCCCTCTACABAATAAATAAAAAAATAAATTT 545
Dd	758 GGAGGAGGGGAGGAGAGAGGCTCTGGGCTCTCTGCAAAAATAAAAAATAAAAAAAT 817
QY	546 TAAATAAT-----AAAAATCCTATATCC-----CATATAAGAAATAAAAGA 586
Dd	818 AAATAAAATTTTAAAAATAATAAAATTCACATATACATATAAGAAATAAAAGAA 877
QY	587 GTCTCAGT----GCAGTATTGGCAAAATTAATCCATTCTTTTTTAATACGGGAATAT-- 640
Dd	878 GTCTCAGTGCAGCATTTTGTCAAAAATTAATATCCATTCTTTTTTATACGGGTGAATAT 937
QY	641 --TGCGCAATTATAGATCTGGATTG-GACCACTTAATGAAGCGGCG-ACCCACAGGTGTTTG 696
Dd	938 TGCGCAATTATAGATCTGGATTTGAACCACTTAATGAAGCGGCGCACACACAGGTGTTTG 997
QY	697 AGGTGTTGGCAATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTAATTAATCTTGCA 756
Dd	998 AGGTGTTGGCAATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTAATTAATCTTGCA 1057
QY	757 AAAATGGTCTGTGCACTTCGAATGTGAAATGCTGTCCAGTTTATTTTTTTTTTATGTTGT 816

Db 1058 AAAAAATGGTCTGTGCACTTGGATGTGAATGCTGCCAGTTTATATTTTATTTATGTTGT 1117
Qy 817 ATCCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTCCTGTTTATTTTGG 876
Db 1118 ATCCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTCCTGTTTATTTGG 1177
Qy 877 TCATCTTTAGAGTTATATCAGGAATGTGTTTAAACAAGAGAAAGAACTTTTCTAAGGAATG 936
Db 1178 TCATCTTTAGAGTTATATCAGGAATGTGTTTAAACAAGAGAGAACTTTTCTAAGGAATG 1237
Qy 937 ATACATAGAAAGATTTTATTTTAAATGAGTGTAAAGCTTGTGTTTCTTTGTTGCTGC 996
Db 1238 ATACATAGAAAGATTTTATTTTAAATGAGTGTAAAGCTTGTGTTTCTTTGTTGCTGC 1297
Qy 997 AAGCTATCTGCCCAAGTTTAAATGCAATGGACACATTTTATGTGAGAAAAACACACACA 1056
Db 1298 AAGCTATCTGCCCAAGTTTAAATGCAATGGACACATTTTATGTGAGAAAAACACACACA 1357
Qy 1057 CAC 1104
Db 1358 CANGNN 1405

RESULT 3
BD276315 1581 bp DNA linear PAT 17-JUL-2003
LOCUS MOLECULES EXPRESSED IN HIPPOCAMPUS.
DEFINITION BD276315
ACCESSION BD276315
VERSION BD276315.1 GI:33086083
KEYWORDS JP 2002543833-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1581)
AUTHORS Baughn,M.R., Lal,P., Yue,H., Tang,T.Y., Azimzai,Y. and Kaser,M.R.
TITLE MOLECULES EXPRESSED IN HIPPOCAMPUS
JOURNAL Patent: JP 2002543833-A 1 24-DEC-2002;
INCYTE GENOMICS INC,Matthew R KASER,Preeti LAL,Henry YUE, Tom Y
TANG, Mariah R BAUGHN,Yalda AZIMZAI
COMMENT OS Homo sapiens
PN JP 2002543833-A/1
PD 24-DEC-2002
PR 10-MAY-2000 JP 2000618442
PR 17-MAY-1999 US 09/313300
PI mariah r baughn,preeti lal,henry yue,tom y tang,yalda azimzai,
PI matthew r kaser
CC This description about <220> can't be interpreted CC <220>
CC <221> misc feature
CC <223> Incyte ID No.: 239240
FH Key Location/Qualifiers.
FEATURES
source
1..1581
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 71.6%; Score 790; DB 6; Length 1581;
Best Local Similarity 90.5%; Pred. No. 2.3e-126;
Matches 985; Conservative 0; Mismatches 70; Indels 33; Gaps 12;
Qy 48 GCGGAGGAGGCGCGCGGAGCCCGGACGCGACCCATGCGGAGGTGCTGCCCTAC 107
Db 1 GCGGAGGAGGCGCGCGGAG-CCCCGACGCGACCATGTCGAGGTGCTGCCCTAC 59
Qy 108 GCGCAGCAGAGCTGAGCCCTACGCGCAGCGCGCGAGCTGGCGCAGATCTTCTCCTGC 167
Db 60 GCGCAGCAGAGCTGAGCCCTACGCGCAGCGCGCGAGCTGGCGCAGATCTTCTCCTGC 119
Qy 168 GCGCTGCGAGGACCAACAACTTCTTTCGCGCGCGGCGAGAAACAAGCGCGCCCAAGCTG 227

Db 120 GCGCTGCGAGGACCAACAACTTCTTTCGCGCGCGGCGAGAAACAAGCGCGCCCAAGCTG 179
Qy 228 GCGCAGATCGGCGGAGCAAGCGGCTTGTATTGAAGATGATAGGATTTGATGACGTGCTG 287
Db 180 GCGCAGATCGGCGGAGCAAGCGGCTTGTATTGAAGATGATAGGATTTGATGACGTGCTG 239
Qy 288 AAAAATATGACCGACAGGCACTT-CTGGTGTCTAACT-CCCCAAAGACAAATGAGTTAAG 345
Db 240 AAAAATATGACCGACAGGCACTTCTGGTGTCTAACTCCCCAAAGACAAATGAGTTAAG 299
Qy 346 GGAGAGAATAGGAACGCGGTAACAGTTATTGGCAAAAGCATGAAAAGAGAAAGCACATT 405
Db 300 GGAGAGAATAGGAACGCGGTAACAGTTATTGGCAAAAGCATGAAAAGAGAAAGCACATT 359
Qy 406 TGAATTTATTACTAGCTTG-TACCACGATGAAATCAACACCTGTATCTGGTAT-ATG 463
Db 360 TGAATTTATTACTAGCTTGCTACCCAGATGAAATCAACACCTGTATCTGGTATCAGG 419
Qy 464 CCGGAGACAGATTAGGCGA-AGGAGGAAGAGAGAGAGAAAGGCTTGGGC-CCTCTA 521
Db 420 CCGGAGACAGATTAGGCGGAGAGGAGGAGGAGGAGGAGGAGGCTCTGGGCTCCTCTG 479
Qy 522 CAAATATAAATAAAAAAATTTTAAATATAAATAATCAAAATCACTATATATACA 539
Db 480 CAAATATAAATAAAAAAATATAAATAATTTTAAAAATATAAATAATCACTATATACA 539
Qy 567 -TCCCATATAAGAATAAAGAGTCTCAGT----GCAGTATTGGCAAAATATAAATCCATT 621
Db 540 CATATAAAGAAATAAAGAGTCTCAGTTCGAGCTATTGTCAGAAATTAATATCCATTT 599
Qy 622 CTTTATAATACGGG----AATATTGGCAATATAGATCTGGATTTTG-ACCACCTTAATGA 676
Db 600 CTTTATAATACGGTGAATATTGGCAATATAGATCTGGATTTTGAAACCACTTAATGA 659
Qy 677 GCGGC-ACCACAGGTGTTTGGGAGTGGCAATCTTTCGCTGATTTGGCTGTTCCTCAATG 735
Db 660 GCGGCAACACAGGTGTTTGGGAGTGGCAATCTTTCGCTGATTTGGCTGTTCCTCAATG 719
Qy 736 TTTACATTTTAAATCTTGCAGAAATGCTTCTGTGCACTTGGATGTGAAATGCTGTCCAG 795
Db 720 TTTACATTTTAAATCTTGCAGAAATGCTTCTGTGCACTTGGATGTGAAATGCTGTCCAG 779
Qy 796 TTTATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGT 855
Db 780 TTTATTTTTTTTATGTTGTTATCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGT 839
Qy 856 AGATATCTGTTTATTTTGGTTCATCTTTAGAGTTATCAGAAATGTTTAAACACAGA 915
Db 840 AGATATCTGTTTATTTTGGTTCATCTTTAGAGTTATCAGAAATGTTTAAACACAGA 899
Qy 916 AGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTATTTTAAATGAGTTGTAAG 975
Db 900 AGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTATTTTAAATGAGTTGTAAG 959
Qy 976 CTTGTGTTTCTTGTGTTGCTGCAAGCTATCTGCCCAAGTTTAAATGCAATGAGACATTTTT 1035
Db 960 CTTGTGTTTCTTGTGTTGCTGCAAGCTATCTGCCCAAGTTTAAATGCAATGAGACATTTTT 1019
Qy 1036 TATGTCAGAAAAACA 1095
Db 1020 TATGTCAGAAAAACA 1079
Qy 1096 AAAAAAA 1103
Db 1080 AAACAAA 1087
RESULT 4
AX049462 1581 bp DNA linear PAT 12-JAN-2001
LOCUS Sequence 1 from Patent WO0070036.
DEFINITION AX049462
ACCESSION AX049462
VERSION AX049462.1 GI:12226181

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Kaser,M.R., Lal,P., Yue,H., Tang,Y.T., Baughn,M.R. and Azimzai,Y.
Genes expressed in hippocampus
Patent: WO 0070036-A 1 23-NOV-2000;
Incyte Genomics, Inc. (US)

FEATURES
source Location/Qualifiers
1..1581
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No.: 239240"

ORIGIN

Query Match 71.6%; Score 750; DB 6; Length 1581;
Best Local Similarity 90.5%; Pred. No. 2.3e-126;
Matches 985; Conservative 0; Mismatches 70; Indels 33; Gaps 12;

Qy 48 GGCGAGGAGCGCGGGCGGAGCCCTCGGACGCACCATGTGGGAGTGTCGCCCTAC 107
Db |
1 GGCGAGGAGCGCGGGCGGAG-CGCCGACGCACCATGTGCGAGTCTGCCCTAC 59

Qy 108 GGCGACGAGACTGAGCCCCTAGCGGACGGCGGACGCTGGCGCAGATCTTCTCTGC 167
Db |
60 GGCGACGAGAAGCTGAGCCCCCTAGCGGACGGCGGACGCTGGCGCAGATCTTCTCTGC 119

Qy 168 CGCTCGAGGACACCAACAATTCTTCGGCGCGGGCGAGAACAAAGCGCGCCCAAGCTG 227
Db |
120 CGCTCGAGGACACCAACAATTCTTCGGCGCGGGCGAGAACAAAGCGCGCCCAAGCTG 179

Qy 228 GGCCAGATCGCCCGGACGAGCGGTTGTTATTGAAGATGATAGATTGATGAGTGGTGTG 287
Db |
180 GGCCAGATCGCCCGGACGAGCGGTTGTTATTGAAGATGATAGATTGATGAGTGGTGTG 239

Qy 288 AAAAATATGACCGACAAGGCACCT-CTGGTGTCTAACT-CCCCAAAGACAATGATTTAAG 345
Db |
240 AAAAATATGACCGACAAGGCACCTCTGGTGTCTAACTCCCCAAAGACAATGATTTAAG 299

Qy 346 GGAGAGAATAGGAACGGCGGTAA CAGTTATTGCAAAGACATGAAAAGAGAAGCACTT 405
Db |
300 GGAGAGAATAGGAACGGCGGTAA CAGTTATTGCAAAGACATGAAAAGAGAAGCACTT 359

Qy 406 TGAATTTATTACTAGCTTG-TACCCACGATGAAATCAACAACCTGTATCTGGTAT-ATG 463
Db |
360 TGAATTTATTACTAGCTTGCTACCACGATGAATCAACAACCTGTATCTGGTATCAGG 419

Qy 464 CC CGGACACAGATTAGCGGA-AGGAGGAAGAGAGAGAAGCAAGGCTTGGGC-CCCTCTA 521
Db |
420 CC CGGACACAGATTAGCGCGAGGAGGAGGAGGAGGAGGAGGCTCTGGGCTCCTCTG 479

Qy 522 CAATAAATAAAAAAATAAAAAATTTAAATAATAAAATCCCTATA-----566
Db |
480 CAATAAATAAAAAATAAAAAATTTAAATAATAAAATTAATAAAATCACTATATACA 539

Qy 567 -TCCATATAAGATAAAAGAGTCTCAGT---GCAGTATTGGCAAAATTAATATCCAATTT 621
Db |
540 CATATAAGAAATAAAAAAGAGTCTCAGTTGCAGTATTGTCAAATTAATATCCAATTT 599

Qy 622 CTTTTTTAACGGG----AAATATGGCAATTATAGATCTGGATTGTC-ACCACTTAATGAA 676
Db |
600 CTTTTTTATACCGTGTTTGAAGTTTGGCAATCTCTCGTGATTTGGCACTTAATGAA 659

Qy 677 GC GGC-ACCCCGAGTGTGTTGAGGTGTGGCATTCTTCGCTGATTTGGCTGTTCCCAATG 735
Db |
660 GC GGCACAACGAGTGTGTTGAGGTGTGGCAATCTCTCGTGATTTGGCACTTAATGAA 719

Qy 736 TTTTACATTTAATCTTGCAAAAATGGTCTCTGTCACCTTGGATGGAATGCTGCCAG 795
Db |
720 TTTTACATTTAATCTTGCAAAAATGGTCTCTGTCACCTTGGATGGAATGCTGCCAG 779

REFERENCE
AUTHORS
TITLE
JOURNAL

1 Plumb, B.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 10, 2000 this sequence version replaced gi:9797484.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dj658J7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 120052 bases at least Q40
Consensus quality: 122543 bases at least Q30
Consensus quality: 123707 bases at least Q20
Insert size: 125041; sum-of-contigs
Insert size: 147762; 1.2% error; agarose-fp
Quality coverage: 3.57x in Q20 bases; sum-of-contigs Quality
coverage: 3.10x in Q20 bases; agarose-fp

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 16101: contig of 16101 bp in length
* 16101: gap of 100 bp
* 16202 19430: contig of 3229 bp in length
* 19431 19530: gap of 100 bp
* 19531 46459: contig of 26929 bp in length
* 46460 46559: gap of 100 bp
* 46560 51855: contig of 5296 bp in length
* 51855: gap of 100 bp
* 51956 58425: contig of 6470 bp in length
* 58426 58525: gap of 100 bp
* 58526 69865: contig of 11340 bp in length
* 69866 69965: gap of 100 bp
* 69966 73734: contig of 3789 bp in length
* 73735 91543: gap of 100 bp
* 91544 91643: contig of 17709 bp in length
* 91644 93828: contig of 2185 bp in length
* 93829 93929: gap of 100 bp
* 93929 102232: contig of 8304 bp in length
* 102233 102332: gap of 100 bp
* 102333 109944: contig of 7612 bp in length
* 109945 110044: gap of 100 bp
* 110045 126141: contig of 16097 bp in length.
Location/Qualifiers

FEATURES
source

1. 126141
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="1"
/map="p36.11-36.13"
/clone="RP4-658J7"
/clone_lib="RPC1-4"
1. 16101

misc_feature

/note="assembly fragment:00015
fragment chain:1"

misc_feature

16202. 19430
/note="assembly fragment:00124
fragment chain:1"

misc_feature

19531. 46459

/note="assembly fragment:00873
fragment chain:1"
46560. 51855
/note="assembly fragment:00305
fragment chain:2"
51956. 58425
/note="assembly fragment:00664
fragment chain:2"
58526. 69865
/note="assembly fragment:00698
fragment chain:3"
69966. 73734
/note="assembly fragment:01008
fragment chain:3"
73835. 91543
/note="assembly fragment:00799
fragment chain:4"
91644. 93828
/note="assembly fragment:00579
fragment chain:4"
93929. 102232
/note="assembly fragment:00217"
102333. 109944
/note="assembly fragment:00284"
110045. 126141
/note="assembly fragment:01133"

ORIGIN

Query Match 55.0%; Score 607.2; DB 2; Length 126141;
Best Local Similarity 89.2%; Pred. No. 1.8e-94;
Matches 787; Conservative 0; Mismatches 63; Indels 32; Gaps 11;
Qy 252 GTTGTATTGAAGATGATAGGATTGATGACGTCTGAAAATATGACCGAAGGCACCT 311
Db 75303 GTTGTATTGAAGATGATAGGATTGATGACGTCTGAAAATATGACCGAAGGCACCT 75362
Qy 312 -CTGGTGTCTAACT-CCCCAAGACAATGAGTTAAGGAGAGAGATAGAAACGGCGGTAAAC 369
Db 75363 CCTGGTGTCTAACTCCCCAAGACAATGAGTTAAGGAGAGAGATAGAAACGGCGGTAAAC 75422
Qy 370 AGTTATTGGCAAAAGCATGAAAAGAGACACTTTGAAATTTATTACTAGCTTG-TAC 428
Db 75423 AGTTATTGGCAAAAGCATGAAAAGAGACACTTTGAAATTTATTACTAGCTTGCTAC 75482
Qy 429 CCAGATGAATCAACCAACCTGTATCTGGTAT-ATGCCCGAGACAGATTAGGCGA-AGG 486
Db 75483 CCAGATGAATCAACCAACCTGTATCTGGTATAGGCCCGGAGACAGATTAGGCGAGAGG 75542
Qy 487 AGGAAGAGAGAGAGAGAGAGAGGCTTGGGC-CCTCTACAAATAAAATAAAAAAATAATT 545
Db 75543 AGGAGGAGGAGGAGAGAGAGGCTCTGGGCTCTCTGCAAAATAAAATAAAAAATAATA 75602
Qy 546 TAAATAATAAATCCCTATA-----TCCCATATAGATATAAAGAGTTC 589
Db 75603 TAAATAATAAATAAATAAATAAATTCATATATACATATAAAGATAAAGAGAGTTC 75662
Qy 590 TCAGT---CGAGTATTGGCAAAATTAATCCATTTCTTTTAAATACGG----AATATT 641
Db 75663 TCAGTTGCAGCTATTGTCAAAATTAATATCCATTTCTTTTATATACGGTGAATATGTC 75722
Qy 642 GGCAATATAGATCTGGATTTTG-ACCATTAAATGAACGGC-ACCCAGGTGTTTTGAGG 699
Db 75723 GCAATATATAGATCTGGATTTTGAACCACTTAATGAACGGCGCAACCCAGGTGTTTTGAGG 75782
Qy 700 TGTGGCATTTCTGCTGATTGTGGCTGTTCCCATGTTTACATTTATTATCTTGCATAA 759
Db 75783 TGTGGCATTTCTGCTGATTGTGGCTGTTCCCATGTTTACATTTATTATCTTGCATAA 75942
Qy 760 ATGTTCTGTGCACCTTGGATGTGAATGCTGTCAGTCTTTTATTTTATTTTATGTTGTTATC 819
Db 75843 ATGTTCTGTGCACCTTGGATGTGAATGCTGTCAGTCTTTTATTTTATTTTATGTTGTTATC 75902
Qy 820 CTTGGATGTACAAAAAATTCAGAAAATGATCTCTGTAGATATTCTCTGTTTATTTTGTGTC 879

[illegible]

RESULT 8	
AL391357/c	
LOCUS	AL391357
DEFINITION	Human DNA sequence from clone RP11-401M16 on chromosome 1, complete sequence.
ACCESSION	AL391357
VERSION	AL391357.20
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 195076)

FEATURES	SOURCE
----------	--------

ORIGIN

	Query Match	55.0%;	Score 607.2;	DB 9;	Length 195076;
	Best Local Similarity	89.2%;	Pred. No. 1.9e-94;		
	Matches 787;	Conservative 0;	Mismatches 63;	Indels 32;	Gaps 11;
Qy	252	GTGTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAATAATATGACCGACAAGGACCT	311		
Db	5527	GTGTGTTATTGAAGATGATAGGATTGATGACGTGCTGAAATAATATGACCGACAAGGACCT	5468		
Qy	312	-CTGGTGTCTAACT-CCCAAGAACAATGAGTTAAGGAGAGAAATAGGAACGCGCGTAACT	369		
Db	5467	CCTGGTGTCTAACTCCCAAGAACAATGAGTTAAGGAGAGAAATAGGAACGCGCGTAACT	5408		
Qy	370	AGTTATTGGCAAAAACGATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTG-TAC	428		
Db	5407	AGTTATTGGCAAAAACGATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTAC	5348		
Qy	429	CCACGATGAAATCAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGCGGA-AGG	486		
Db	5347	CCACGATGAAATCAACAACCTGTATCTGGTATCAGCCCGGAGACAGATTAGCGCGAGAGG	5288		
Qy	487	AGGAAGAGAGAGAGAAGAAGCGCTCGGC-CCTCTACAAATAAAATAAAAAAAAATTT	545		
Db	5287	AGGAAGAGAGAGAGAAGGCTCTGGGCTCTCTGCATAAATAAATAAATAAATAAATAA	5228		
Qy	546	TAAATTAATAAATCCCTATA-----TCCCATATAGAAATAAAGAGATC	589		
Db	5227	TAAATTTTAAAAAATAATAAATAATTCACATATACACATATAAAGAAATAAAGAAAGTCT	5168		
Qy	590	TCAGT----GCAGTATTGCAAAATAAATCCATTTCTTTTAAATACGG----ATAATT	641		
Db	5167	TCAGTTGCGAGCTATTGTCAAATAATATATCAATTTCTTTTATATACGGTGAATATTGC	5108		
Qy	642	GGCATTTATAGATCTGGAATTTTG-ACCACATTAATGAAGCGGC-ACCCAGAGTGTTTTGAGG	699		
Db	5107	GCAATTTATAGATCTGGAATTTTGAACCCACATTAATGAAGCGCAACACACAGGTGTTTGAGG	5048		
Qy	700	TGTTGGCATTTCTGGCTGATTGGCTGTTCGCCAATGTTTACATTTATTAATCTTGCAAAA	759		
Db	5047	TGTTGGCATTTCTGGCTGATTGGCTGTTCGCCAATGTTTACATTTATTAATCTTGCAAAA	4988		
Qy	760	ATGGTTCTGTGCACCTGGAGTGAATAATGCTGTCCAGTTTTATTTTTTATATGTGTATC	819		
Db	4987	ATGGTTCTGTGCACCTGGAGTGAATAATGCTGTCCAGTTTTATTTTTTATGTGTATC	4928		
Qy	820	CTTGGATGTACAAAAAATTACAAAAATGATCTCTGTAGATATCTCTGTTTTATTTTGGTCA	879		
Db	4927	CTTGGATGTACAAAAAATTACAAAAATGATCTCTGTAGATATCTCTGTTTTATTTTGGTCA	4868		

Db 541 TCGAATAATGGTCTGTGTCACCTGGATGTGAATGCTGTCCAGTTTTATTTTTTTATGT 600

QY 813 TGTATCTCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATTCGTGTTTAT 872

Db 601 TGTATCTCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATTCGTGTTTAT 660

QY 873 TTGGTCATCTTTAGAGTTATCAGGAATGTGTTTAAACAAGAGAACTTTCTTAAGG 932

Db 661 TTGGTCATCTTTAGAGTTATCAGGAATGTGTTTAAACAAGAGAACTTTCTTAAGG 720

QY 933 AATGATACATAGAAAGATTTTATTTTAAATGAGTTGTAAGCTTGTTCTTTGTTG 992

Db 721 AATGATACATAGAAAGATTTTATTTTAAATGAGTTGTAAGCTTGTTCTTTGTTG 780

QY 993 CTCGAAGCTATCTGCCAAGTTAATGCAATGACACATTTTTTTATGTGAGAAAAACACA 1052

Db 781 CTCGAAGCTATCTGCCAAGTTAATGCAATGACACATTTTTTTATGTGAGAAAAACACA 840

QY 1053 CAC 1103

Db 841 CAC 891

RESULT 10

AC025224

LOCUS

DEFINITION Homo sapiens chromosome 1 clone RP11-4L9, WORKING DRAFT SEQUENCE, 31 unordered pieces.

ACCESSION AC025224

VERSION AC025224.3 GI:7596970

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 163801)

Waterston,R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 163801)

Waterston,R.H.

Direct Submission

Submitted (07-MAR-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA

On Apr 19, 2000 this sequence version replaced gi:7321912.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

----- Project Information -----

Center project name: H_NH0004109

----- Summary Statistics -----

Sequencing vector: pBluescript

Sequencing strategy: 100% of reads

Chemistry: Dye-terminator Big Dye, 0% of reads

Assembly program: Phrap, version 0.990319

Consensus quality: 145820 bases at least Q40

Consensus quality: 150999 bases at least Q30

Consensus quality: 153311 bases at least Q20

Insert size: 185000; agarose-gel

Insert size: 160801; sum-of-contigs

Quality coverage: 3.57 in Q20 bases; agarose-gel

Quality coverage: 4.25 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence

FEATURES

source

1..163801

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP11-4L9"

1..1640

misc_feature

/note="assembly_name:Contig12"

1741..3443

misc_feature

* as soon as it is available and the accession number will be preserved.

1 1640: contig of 1640 bp in length

1740: gap of unknown length

1741 3443: contig of 1703 bp in length

3444 3543: gap of unknown length

3544 4790: contig of 1247 bp in length

4791 4890: gap of unknown length

4891 6777: contig of 1887 bp in length

6778 8114: contig of 1237 bp in length

8115 8214: gap of unknown length

8215 9912: contig of 1698 bp in length

9913 10012: gap of unknown length

10013 11754: contig of 1742 bp in length

11755 11854: gap of unknown length

11855 14058: contig of 2204 bp in length

14059 14158: gap of unknown length

14159 18450: contig of 4292 bp in length

18451 18550: gap of unknown length

18551 22470: contig of 3920 bp in length

22471 22570: gap of unknown length

22571 26337: contig of 3767 bp in length

26338 26437: gap of unknown length

26438 30118: contig of 3681 bp in length

30119 30218: gap of unknown length

30219 33161: contig of 2943 bp in length

33162 33261: gap of unknown length

33262 37679: contig of 4418 bp in length

37680 37779: gap of unknown length

37780 41155: contig of 3376 bp in length

41156 41255: gap of unknown length

41256 45149: contig of 3894 bp in length

45150 45249: gap of unknown length

45250 50155: contig of 4906 bp in length

50156 50255: gap of unknown length

50256 53871: contig of 3616 bp in length

53872 53971: gap of unknown length

53972 58814: contig of 4843 bp in length

58815 58914: gap of unknown length

58915 63420: contig of 4506 bp in length

63421 63520: gap of unknown length

63521 67773: contig of 4253 bp in length

67774 73896: contig of 6023 bp in length

73897 73996: gap of unknown length

73997 79302: contig of 5306 bp in length

79303 79402: gap of unknown length

79403 85385: contig of 5983 bp in length

85386 85485: gap of unknown length

85486 93969: contig of 8484 bp in length

93970 94069: gap of unknown length

94070 102628: contig of 8559 bp in length

102629 102728: gap of unknown length

102729 113357: contig of 10629 bp in length

113358 113457: gap of unknown length

113458 123217: contig of 9760 bp in length

123218 123317: gap of unknown length

123318 132911: contig of 9594 bp in length

132912 133011: gap of unknown length

133012 146300: contig of 13289 bp in length

146301 146401: gap of unknown length

146402 163801: contig of 17401 bp in length.

PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40

CC Key Location/Qualifiers.

FEAT

Source

1. .524
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 38.7%; Score 427.2; DB 6; Length 524;

Best Local Similarity 97.8%; Pred. No. 8e-64; Indels 2; Gaps 2;

Matches 451; Conservative 2; Mismatches 6; Indels 2; Gaps 2;

645 ATTATAGATCTGATTTTG-ACCACTTAATGAAGCGC-ACCCAGGTTTGTGAGGTGT 702

24 ATTATAGATCTGATTTTGAAACCACTTAATGAAGCGCAACACAGGTGTTTGAGGTGT 83

703 TGGCATTCTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTTGCAGAAATG 762

84 TGGCATTCTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTTGCAGAAATG 143

763 GTTCTGTGCACCTTGGATGTAATGCTGCTCCAGTTTATTTTATGTTGTTATCCCTT 822

144 GTTCTGTGCACCTTGGATGTAATGCTGCTCCAGTTTATTTTATGTTGTTATCCCTT 203

823 GAGTGTACAAAAATTCAGAAATGATCTCTGTAGATATCTGTTTTATTTTGGTCATCT 882

204 GAGTGTACAAAAATTCAGAAATGATCTCTGTAGATATCTGTTTTATTTTGGTCATCT 263

883 TTAGAAGTTATCAGGAATGTTTAAACAAGAGAGAACTTTCTAAGGAATGATACAT 942

264 TTAGAAGTTATCAGGAATGTTTAAACAAGAGAGAACTTTCTAAGGAATGATACAT 323

943 AGAAAAAGATTTATTTTAAAAAGAGTTGTAAGAGCTTGTTCTTTTGTGCTCAAGCTA 1002

324 AGAAAAAGATTTATTTTAAAAAGAGTTGTAAGAGCTTGTTCTTTTGTGCTCAAGCTA 383

1003 TCTGCCCAAGTTATGCCAAATGACACATTTTATGTAGCAAAAAACACACACACACA 1062

384 TCTGCCCAAGTTATGCCAAATGACACATTTTATGTAGCAAAAAACACACACACACA 443

1063 CACACACACACACACACACAGAAAAAAGAAAAA 1103

444 CACACACACACACACACACAGAAAAAAGAAAAA 484

RESULT 14

BD139338

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD139338 417 bp DNA linear PAT 18-SEP-2002

Extended cDNA of secretory protein.

BD139338

BD139338.1 GI:23234283

JP 2002508182-A/90.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 417)

Bougueleret, L., Duclert, A. and Edwards, J.B.D.M.

Patent: JP 2002508182-A 90 19-MAR-2002;

GENSET

OS Homo sapiens (human)

PN JP 2002508182-A/90

PD 19-MAR-2002

PR 17-DEC-1998 JP 2000539136

PR 17-DEC-1997 US 60/069957,09-FEB-1998 US 60/074121 PR

13-APR-1998 US 60/081563,10-AUG-1998 US 60/096116 PI

BOUGUELERET,AYMERIC DUCLERT,JEAN BAPTISTE DUMAS MILNE PI EDWARDS

PC C12N15/09,C12N15/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
C12N1/21,

PC C12N5/10,C12P21/02,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC

Extended cDNA of secretory protein

FEAT Key Location/Qualifiers

FT CDS 327..416

FT polyA site 404..417.

Location/Qualifiers

1..417

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

Query Match 36.7%; Score 405.4; DB 6; Length 417;

Best Local Similarity 98.6%; Pred. No. 4.5e-60;

Matches 409; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

690 TGTTTTGAAGGTGTTGGCAATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTATAA 749

1 TGTTTTGAAGGTGTTGGCAATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTATAA 60

750 TCTTGCAAAAAATGTTCTGTGCACCTTGGAATGTGAAAATGCTGCCAGTTTATTTTATA 809

61 TCTTGCAAAAAATGTTCTGTGCACCTTGGAATGTGAAAATGCTGCCAGTTTATTTTATA 120

810 TGTGTTTATCCTCGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATCTGTTTT 869

121 TGTGTTTATCCTCGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATCTGTTTT 180

870 ATTTTGGTCATCTTTTAGAAGTTATCAGGAATGTGTTTAAAAACAAGAGAACCTTTTCTA 929

181 ATTTTGGTCATCTTTTAGAAGTTATCAGGAATGTGTTTAAAAACAAGAGAACCTTTTCTA 240

930 AGGAATGATACATAGAAAAAGATTTTATTTTAAAAATGAGTTGTAAGCTGTGTTCTTTG 989

241 AGGAATGATACATAGAAAAAGATTTTATTTTAAAAATGAGTTGTAAGCTGTGTTCTTTG 300

990 TTGCTGCAAGCTATCTGCCCAAGTTTAAATGCAAAATGGACACATTTTATGTCAGAAAAAC 1049

301 TTGCTGCAAGCTATCTGCCCAAGTTTAAATGCAAAATGGACACATTTTATGTCAGAAAAAC 360

1050 AC 1104

361 AC 415

RESULT 15

CQ704178

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..366

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

```
Qy 731 CAATGTTTACATTATTATCTTGCAAAAATGGTTCTGTGCACCTTGGATGTGAAATGCTG 790
Db |||||
Qy 1 CAATGTTTACATTATTATCTTGCAAAAATGGCTCTGTGCACCTTGGATGTGAAATGCTG 60
Db |||||
Qy 791 TCCAGTTTTATTTTATGTTTATGTTTATCCCTTGGATGTACAAAAAATTCAGAAAATGATC 850
Db |||||
Qy 61 TCCAGTTTTATTTTATGTTTATGTTTATCCCTTGGATGTACAAAAAATTCAGAAAATGATC 120
Db |||||
Qy 851 TCTGTAGATATTCTGTTTATTTTATTTTGGTCACTTTTAGAAGTTATCAGGAATGTGTTTAAAA 910
Db |||||
Qy 121 TCTGTAGATATTCTGTTTATTTTATTTTGG-CATCTTTAGAAGTTATCAGGAATGTGTTTAAAA 179
Db |||||
Qy 911 CAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTATTTTAAAAATCAGTTG 970
Db |||||
Qy 180 CAAGAAGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTATTTTAAAAATCAGTTG 239
Db |||||
Qy 971 TAAAGCTTGTGTTTCTTGTGCTGCAAGCTATCTGCCCAAGTTAATGCAATGGACACA 1030
Db |||||
Qy 240 TAAAGCTTGTGTTTCTTGTGCTGCAAGCTATCTGCCCAAGTTAATGCAATGGACACA 299
Db |||||
Qy 1031 TTTTATGTTCAGAAAAACACACACACACACACACACACACACACACACACACACACACAC 1090
Db |||||
Qy 300 TTTTATGTTCAGAAAAACACACACACACACACACACACACACACACACACACACACACAC 359
Db |||||
Qy 1091 AAAAAA 1097
Db |||||
Qy 360 AAAACA 366
```

Search completed: March 30, 2005, 02:07:57
Job time : 4821 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2005, 17:36:37 ; Search time 633 Seconds
(without alignments)
10324.478 Million cell updates/sec

Title: US-10-071-645-1
Perfect score: 1104
Sequence: 1 ggcacgagcgagccgcgtg.....aaaaaaaaaaaaaaaaaaaa 1104

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1104	100.0	1104	AD31052	Ad31052 Human PAP
2	803.8	72.8	1876	AD161702	Ad161702 Human CDN
3	790	71.6	1581	AAC87358	Aac87358 Human dev
4	699.8	63.4	1510	AX20426	Ax20426 Human sec
5	699.8	63.4	1510	ADD90210	Add90210 Novel hum
6	699.8	63.4	1510	ADG90029	Adg90029 Human CDN
7	642	58.2	2058	ADQ24238	Adq24238 Human sof
8	634.6	57.5	1850	ACN90200	Acn90200 Breast ca
9	632.6	57.3	1321	ADJ75198	Adj75198 Marker ge
10	632.6	57.3	1321	ADN04817	Adn04817 Antipsori
11	632.6	57.3	1321	ADR25935	Adr25935 Breast ca
12	607.2	55.0	3189	AD31053	Ad31053 Human PAP
13	519.8	47.1	873	ABQ60975	Abq60975 PRO1489 p
14	443.2	40.1	553	ABX97604	Abx97604 cDNA enco
15	427.2	38.7	524	AC06588	Aac06588 Human sec
16	405.4	36.7	417	ADP18899	Adp18899 Human sec
17	403.8	35.6	417	AX97632	Aax97632 Extended
18	282.2	25.6	411	ABL62340	Ab162340 Colon ade
19	282.2	25.6	411	ABL65003	Ab165003 Lung canc
20	282.2	25.6	411	ABL67288	Ab167288 Thyroid c

21	282.2	25.6	411	6	ABL66839	Ab166839 Lung canc
22	272	24.6	620	11	ACN87560	Acn87560 Breast ca
c 23	270	24.5	417	5	AAF68540	Aaf68540 Human lun
c 24	270	24.5	417	6	ABK38451	Abk38451 cDNA enco
c 25	270	24.5	417	8	ACA10780	Aca10780 Human lun
c 26	270	24.5	417	8	ABX99731	Abx99731 Lung canc
c 27	270	24.5	417	10	ADH45994	Adh45994 Human lun
c 28	270	24.5	417	12	ADE71746	Ade71746 Human lun
c 29	270	24.5	417	13	ADJ19913	Adj19913 Human lun
c 30	262	23.7	453	6	ABV97214	Abv97214 Human pan
c 31	259.4	23.5	451	11	ADT96448	Adt96448 Colon can
c 32	235.6	20.4	332	2	AAT19130	Aat19130 Human gen
c 33	197.8	17.9	549	10	ABT40422	Abt40422 Toxicity
c 34	197.8	17.9	549	12	ADP71513	Adp71513 Renal tox
c 35	185.8	16.8	299	10	ADL24606	Adl24606 Intestina
36	165	14.9	336	4	AAD23580	Aad23580 Human lun
37	165	14.9	336	10	ADD66868	Add66868 Human lun
38	165	14.9	336	10	ADE88122	Ade88122 Human lun
c 39	162.4	14.7	869	6	ABQ43208	Abq43208 Oligonucl
40	162.4	14.7	869	6	ABQ43209	Abq43209 Oligonucl
c 41	155.8	14.1	183	3	AAC06461	Aac06461 Human sec
c 42	149.8	13.6	869	6	ABQ43206	Abq43206 Oligonucl
c 43	149.8	13.6	869	6	ABQ43207	Abq43207 Oligonucl
44	146.2	13.2	654	13	ADQ52714	Adq52714 Novel can
45	144	13.0	329	2	AAQ60934	Aaq60934 Human bra

ALIGNMENTS

RESULT 1
AAD31052
ID AAD31052 standard; cDNA; 1104 BP.

AC AAD31052;

DT 18-JUN-2002 (first entry)

DE Human PAPAP cDNA.

KW Human; PAPAP protein; schizophrenia candidate gene; g34872 gene;
schizophrenia; bipolar disorder; central nervous system disorder;
psychotic disorder; mood disorder; autism; mental retardation;
psychiatric disease; anxiety disorder; impulse-control disorder;
eating disorder; cognitive disorder; personality disorder; vaccine;
chromosome 1p35-p36; neuroleptic; antialcoholic; tranquiliser;
antidepressant; nootropic; antiaddictive; ss.

OS Homo sapiens.

Key Location/Qualifiers
5'UTR 1..86
CDS 87..344
FT /*tag= a
FT /*tag= b
FT /*product= "Human PAPAP protein"
FT 345..1104
FT /*tag= c
FT polyA_site 1085..1104
FT /*tag= d

WO200212279-A2.

14-FEB-2002.

26-JUL-2001; 2001WO-IB001891.

07-AUG-2000; 2000US-0223482P.

(GEST) GENSET.

PI Bihaïn B, Bour B, Bougueleret L;

XX

DR WP1; 2002-241732/29.
XX P-PSDB; AAE19431.
PT Novel isolated and purified or recombinant polynucleotide encoding PAPAP
PT protein, useful for diagnosing and treating schizophrenia, bipolar
XX disorder and other central nervous system disorders.
PS
XX Claim 2; Page 92; 96pp; English.
CC The invention relates to human PAPAP polypeptides and polynucleotides.
CC The invention also concerns the interaction of PAPAP with schizophrenia
CC candidate gene g34872. PAPAP polypeptides, gene and anti-PAPAP antibodies
CC are useful for treating schizophrenia, bipolar disorder or related
CC central nervous system (CNS) disorders e.g. psychotic disorders, mood
CC disorders, autism, substance dependence and alcoholism, mental
CC retardation and other psychiatric diseases including cognitive, anxiety,
CC eating, impulse-control and personality disorders. They are also used in
CC vaccines. The present sequence is human PAPAP cDNA. The PAPAP gene is
XX located on chromosome 1p35-p36
SQ Sequence 1104 bp; 350 A; 220 C; 262 G; 272 T; 0 U; 0 Other;
Query Match 100.0%; Score 1104; DB 6; Length 1104;
Best Local Similarity 100.0%; Pred. No. 3.4e-219;
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAACGAGGCGCGCGCTGACCTCTCGCGCGCGCGGACGCGCGGCGGAGGCGG 60
DB 1 GGCAACGAGGCGCGCGCTGACCTCTCGCGCGCGCGGACGCGCGGCGGAGGCGG 60
QY 61 CGCGCGCGGAGCCCCCGGACCGACCATGTGGAGGTGCTGCCCTACGGCGACGAGGC 120
DB 61 CGCGCGCGGAGCCCCCGGACCGACCATGTGGAGGTGCTGCCCTACGGCGACGAGGC 120
QY 121 TGAGCCCTACGGCGAGCGGCGGACGTGGCGCAGATCTTCTCGCCGCTCGAGGACA 180
DB 121 TGAGCCCTACGGCGAGCGGCGGACGTGGCGCAGATCTTCTCGCCGCTCGAGGACA 180
QY 181 CCAACAACTCTTTCGGCGCGCGGACGAGCAAGCGCGCGCCCAAGCTGGCGCAGATCGGC 240
DB 181 CCAACAACTCTTTCGGCGCGCGGACGAGCAAGCGCGCGCCCAAGCTGGCGCAGATCGGC 240
QY 241 GGAGCAAGCGGCTGTGTTATGAGATGATGAGATGATGATGATGATGATGATGATGATG 300
DB 241 GGAGCAAGCGGCTGTGTTATGAGATGATGAGATGATGATGATGATGATGATGATGATG 300
QY 301 ACAAGGCACTCTGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 360
DB 301 ACAAGGCACTCTGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 360
QY 361 GCGGTAAACAGATTATTTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTTACTA 420
DB 361 GCGGTAAACAGATTATTTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTTACTA 420
QY 421 GCTTTGACCAAGATGAAATCAACACTGTATCTGTTATATGCGCGGAGACAGATTAGG 480
DB 421 GCTTTGACCAAGATGAAATCAACACTGTATCTGTTATATGCGCGGAGACAGATTAGG 480
QY 481 CGAAGGAGGAG 540
DB 481 CGAAGGAGGAG 540
QY 541 AAATTTAAATATAAATCCCTATATCCATATAGAAATAAAGAGTCTCAGTGCAGTA 600
DB 541 AAATTTAAATATAAATCCCTATATCCATATAGAAATAAAGAGTCTCAGTGCAGTA 600
QY 601 TTGGCAAAATTAATCCATTTCTTTTAAATACGGGAATATGGCAATATAGATCTGGATT 660
DB 601 TTGGCAAAATTAATCCATTTCTTTTAAATACGGGAATATGGCAATATAGATCTGGATT 660
QY 661 TTGACCACTTAATGAAGCGGACCCCGAGGTGTTTGGAGTGTGGCATCTTCGCTGATT 720
DB 661 TTGACCACTTAATGAAGCGGACCCCGAGGTGTTTGGAGTGTGGCATCTTCGCTGATT 720

QY 721 TGGCTGTTCCCAATGTTTACATTTATTAATCTTGCABAAATGGTTCTGTGCACTTGGATG 780
DB 721 TGGCTGTTCCCAATGTTTACATTTATTAATCTTGCABAAATGGTTCTGTGCACTTGGATG 780
QY 781 TGAATATGCTGTCAGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 840
DB 781 TGAATATGCTGTCAGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 840
QY 841 GAAATATGCTGCTGATATTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 900
DB 841 GAAATATGCTGCTGATATTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 900
QY 901 GTGTTTAAACAAGAGAGAACTTTCTTAAGGAATGATACATAGAAAAGATTTTATTTA 960
DB 901 GTGTTTAAACAAGAGAGAACTTTCTTAAGGAATGATACATAGAAAAGATTTTATTTA 960
QY 961 AAATGAGTTGTAAGCTTTGTTTCTTTTGTGTCGAAGCTATCTGCCCAAGTTAATGCA 1020
DB 961 AAATGAGTTGTAAGCTTTGTTTCTTTTGTGTCGAAGCTATCTGCCCAAGTTAATGCA 1020
QY 1021 AATGACACATTTTATGTCAGAAAAACACACACACACACACACACACACACACACACA 1080
DB 1021 AATGACACATTTTATGTCAGAAAAACACACACACACACACACACACACACACACACA 1080
QY 1081 CACGAAAAAAGAAAAA 1104
DB 1081 CACGAAAAAAGAAAAA 1104
RESULT 2
ADI61702
ID ADI61702 standard; cDNA; 1876 BP.
XX AC ADI61702;
XX DT 22-APR-2004 (first entry)
XX DE Human cDNA downregulated in Alzheimer's disease, INCYTE 234151.1.
XX KW Human; ss; Alzheimer's disease; differential display; neuroprotective;
XX OS Homo sapiens.
XX PN US6682888-B1.
XX PD 27-JAN-2004.
XX PF 05-MAY-2000; 2000US-00566921.
XX PR 05-MAY-2000; 2000US-00566921.
XX PA (INCY-) INCYTE CORP.
XX PI Loxing JF, Tingley DW, Edwards CM;
XX DR WPI; 2004-118572/12.
XX PT New composition comprising cDNAs that are differentially expressed in
XX brain disorders, useful for diagnosing or treating Alzheimer's disease.
XX PS Claim 1; SEQ ID NO 70; 223pp; English.
XX CC The invention relates to a new composition comprising ADI61633-
XX ADI61702 and their complements that are cDNAs differentially expressed in
XX brain disorders. Also included are a high throughput method for detecting
XX differential expression of one or more cDNAs in a sample containing
XX nucleic acids and a high throughput method for screening a library of
XX molecules or compounds to identify a ligand that specifically binds a
XX cDNA. The expression of the each of the cDNAs is downregulated at least
XX two-fold in the brain of the subjects with Alzheimer's disease (ADI61633-
XX ADI61727) or upregulated at least two fold in Alzheimer's disease

CC (AD161728-ADI61770). The composition is useful for diagnosing or treating
CC Alzheimer's disease. The present sequence is a cDNA downregulated at
CC least two-fold in the brain of the subjects with Alzheimer's disease.

XX
SQ Sequence 1876 BP; 480 A; 442 C; 442 G; 480 T; 0 U; 32 Other;
Query Match 72.8%; Score 803.8; DB 12; Length 1876;
Best Local Similarity 87.9%; Pred. No. 7.7e-157;
Matches 991; Conservative 0; Mismatches 103; Indels 34; Gaps 11;

QY 10 CAGCCGCCGTGACCTGTCGCGCGGGGCGGAGCGGGCGGAGGAGCGCGCGCGG 69
DB |||||
279 CAGCCGCCGTGACCTGTCGCGCGGGGCGGAGCGGGCGGAGGAGCGCGCGCGG 338
QY 70 AGCCCCCGGACGCGACATGTGGGAGGTCTGCCCTACGGCGACGAGAGCTGAGCCCT 129
DB |||||
339 AG-CCCCCGGACGCGACATGTGGGAGGTCTGCCCTACGGCGACGAGAGCTGAGCCCT 397
QY 130 ACGGCGAGCGGCGGACGCTGGGCGAGATCTTCTCTGCGGCTGAGGACCAACAAT 189
DB |||||
398 ACGGCGAGCGGCGGACGCTGGGCGAGATCTTCTCTGCGGCTGAGGACCAACAAT 457
QY 190 TCTTCGGCGCGGCGGACGCGCGCGCCCAAGCTGGGCCAGATCGGCGGCGCAAGC 249
DB |||||
458 TCTTCGGCGCGGCGGACGCGCGCGCCCAAGCTGGGCCAGATCGGCGGCGCAAGC 517
QY 250 GGGTGTGTTTGAAGATGATAGGATGATGAGCTGCTGAAATATGACCGCAAGGCAC 309
DB |||||
518 GGGTGTGTTTGAAGATGATAGGATGATGAGCTGCTGAAATATGACCGCAAGGCAC 577
QY 310 CT-CTGGTGTCTAACT-CCCCAAGACAAATGATTAAGGAGAGAAATAGGAACGGCGTA 367
DB |||||
578 CTCTCTGGTGTCTAACTCCCCAAGACAAATGATTAAGGAGAGAAATAGGAACGGCGTA 637
QY 368 ACAGTATTGGCAAAAGCATGAAAGAGAGAGACACTTTGAAATTTATCTAGCTTG-T 426
DB |||||
638 ACAGTATTGGCAAAAGCATGAAAGAGAGAGACACTTTGAAATTTATCTAGCTTGCT 697
QY 427 ACCCAGCATGAATCAACAACTGTATCTGTGTAT-ATGCCCGGACAGATGAGCGGAAG 485
DB |||||
698 ACCCAGCATGAATCAACAACTGTATCTGTGTATCAGGCGGCGGACAGATGAGCGGAGA 757
QY 486 GAGGAGAGAGAGAGAGAGAGGCTTGGGCGCTCTACAAATTAATAAAAAAATTT 545
DB |||||
758 GGAGGAGGAGGAGGAGAGAGGCTCTGGGCTCTCTGCAAAATAAAAAAATAAAAAAT 817
QY 546 TAAATATAT-----AAATCCCTATATCC-----CATATAGATTAAGA 586
DB |||||
818 AAATATAATTTTAAATAATTAATAATTCACATATATACATATTAAGAATAAAAAAGAA 877
QY 587 GTCTCAGT---GCAGTATTGGCAAAATTAATCCATTTCTTTTAAATACGGGAATAT-- 640
DB |||||
878 GTCTCAGTTCGAGCTATTGTCAAAATTAATATCAATTTCTTTTATATACGGTGAATAT 937
QY 641 --TGCATATATAGATCTGATTTTG-ACACTTAATGAAGCGG-ACCCAGGTTTGTG 696
DB |||||
938 TCGCAATATATAGATCTGATTTTGAACCACTTAATGAAGCGGCAACACAGGTTTGTG 997
QY 697 AGGTGTGGCATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTTAAATCTTCCA 756
DB |||||
998 AGGTGTGGCATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTTAAATCTTCCA 1057
QY 757 AAATGGTTCGTGACATTTGATGGAATGCTGTCAGTGTATTATTTTATATCTGTGT 816
DB |||||
1058 AAATGGTTCGTGACATTTGATGGAATGCTGTCAGTGTATTATTTTATATCTGTGT 1117
QY 817 ATCTTTGATGTACAAAAATTCAGAAATGATCTCTGATGATATCTGTTTTATTTTGG 876
DB |||||
1118 ATCTTTGATGTACAAAAATTCAGAAATGATCTCTGATGATATCTGTTTTATTTTGG 1177
QY 877 TCATCTTTAGATTTATCAGGATGTTGTTTAAACAAGAGACATTTTCTAAGGATG 936
DB |||||
1178 TCATCTTTAGATTTATCAGGATGTTGTTTAAACAAGAGACATTTTCTAAGGATG 1237

QY 937 ATACATAGAAAGATTTTATTTTAAATGAGTTGTAAGCTTGTGTTCTTTTGTCTGC 996
DB |||||
1238 ATACATAGAAAGATTTTATTTTAAATGAGTTGTAAGCTTGTGTTCTTTTGTCTGC 1297
QY 997 AAGCTATCTGCCCAAGTTAATGCAATGGACACATTTTATGTCAGAAAAACACACACA 1056
DB |||||
1298 AAGCTATCTGCCCAAGTTAATGCAATGGACACATTTTATGTCAGAAAAACACACACA 1357
QY 1057 CAC 1104
DB |||||
1358 CANGNN 1405

RESULT 3

AAC87358

ID AAC87358 standard; cDNA; 1581 BP.

XX AAC87358;

XX 09-MAR-2001 (first entry)

XX Human developmentally-regulated hippocampus EST, SEQ ID NO:1.

XX Human hippocampal gene; hippocampus EST; expressed sequence tag;
KW developmental regulation; drug screening; modulator;
KW expression alteration; Alzheimer's disease; Huntington's disease;
KW schizophrenia; epilepsy; diagnostic probe; gene therapy; neurotropic;
KW neuroprotective; anticonvulsant; neuroleptic; ss.

OS Homo sapiens.

XX WO200070036-A2.

XX 23-NOV-2000.

XX 10-MAY-2000; 2000MO-US013046.

XX 17-MAY-1999; 99US-00313300.

XX (INCY-) INCYTE GENOMICS INC.

XX Kaser MR, Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y;

XX WPI; 2001-016229/02.

XX Nucleic acids expressed in hippocampus useful for diagnosing, treating or
PT preventing diseases associated with the hippocampus, e.g., Alzheimer's
PT disease, Huntington's disease, Schizophrenia and epilepsy.

XX Claim 2; Page 31; 39pp; English.

XX The invention relates to novel human developmentally-regulated genes
CC which are expressed in the hippocampus (ESTs given in AAC87358-C87364),
CC and to a protein product of a gene of the invention (AAB29800). The human
CC developmentally-regulated hippocampal gene ESTs (expressed sequence tags)
CC were identified on the basis of homology with ESTs (AAC87365-C87371)
CC found in adult rat hippocampus but not in foetal rat hippocampus. The
CC invention also relates to expression vectors and host cells comprising a
CC human developmentally-regulated hippocampal gene; recombinant production
CC of the gene product; and a method of screening for a compound which
CC specifically binds to a human hippocampal gene of the invention or its
CC product. It additionally encompasses nucleic acid sequences with at least
CC 70% identity to one of the human hippocampal ESTs, and proteins with at
CC least 85% identity to the human hippocampal protein of the invention. The
CC human hippocampal genes, their encoded proteins and modulators of the
CC genes or proteins are useful for treating or preventing diseases
CC associated with altered expression of a gene in the hippocampus, such as
CC Alzheimer's disease, Huntington's disease, schizophrenia, epilepsy and
CC complications of these conditions. The nucleic acid sequences, or
CC fragments thereof are also as hybridisation probes for use in diagnosing
CC a disease or condition associated with altered expression of a
CC hippocampal gene. The present sequence represents a human hippocampal EST

XX This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number is given in the descriptor line. The gene
CC can be used to generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. AAX20403) for increasing the stability of
CC the fused protein as compared to the human protein only. The invention
CC relates to 83 novel genes and their fragments (nucleic acid sequences:
CC AAX20412-X20499; amino acid sequences AAY00258-Y00377) which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. Also, pathological conditions can be diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on
CC which tissues they are most highly expressed in (see AAX20412 for
CC described uses)

XX SQ Sequence 1510 BP; 473 A; 285 C; 299 G; 450 T; 0 U; 3 Other;

Query Match 63.4%; Score 699.8; DB 2; Length 1510;
Best Local Similarity 89.7%; Pred. No. 2.7e-135;
Matches 893; Conservative 1; Mismatches 69; Indels 33; Gaps 12;
QY 141 GCGCAGCGGGCCGACATCTTCTCCTCGCGCTGCGAGCACCAACAATCTTTCGCGGCC 200
DB 2 GCGCAGCGGGCCGACATCTTCTCCTCGCGCTGCGAGCACCAACAATCTTTCGCGGCC 61
QY 201 GCGCAGAACAGCGCGCCCAAGCTGCGCGCAGATCGCGCGAGCAAGCGGTTGTATT 260
DB 62 GCGCAGAACAGCGCGCG-CAAGCTGCGCGAGATCGCGCGAGCAAGCGGTTGTATT 120
QY 261 GAAGATGATAGGATTGATGACGTGCTGTAATAATATGACCGCAAGGACCT-CTGGTGT 319
DB 121 GAAGATGATAGGATTGATGACGTGCTGTAATAATATGACCGCAAGGACCTCTCGTGT 180
QY 320 TAACT-CCCCAAGCAATGAGTAAAGGAGAGATAGCAAGCGGTAACAGTTATTGG 378
DB 181 TAACTCCCCAAGCAATGAGTAAAGGAGAGATAGCAAGCGGTAACAGTTATTGG 240
QY 379 CAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTG-TACCCACGATGA 437
DB 241 CAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCACGATGA 300
QY 438 AATCAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGGCGA-AGGAGGAAGAGA 495
DB 301 AATCAACAACCTGTATCTGGTATCAGCGCGGAGACAGATAGGCGGAGGAGGAGGAGG 360
QY 496 GAGGAGAGAGAGCTTGGC-CTCTCAATAATAATAATAATAATAATAATAATAATAAT 554
DB 361 AGGAGGAGAGAGCTTGGGCTCTCTGCAATAATAATAATAATAATAATAATAATAATTT 420
QY 555 AAAATCCCTTATA-----TCCCATATAAGAAATAAAAGAGTCTCAGT---- 594
DB 421 AAAATATAAAAATTCATATATACATATATAAGAAATAATAAGAGTCTCAGTTGCA 480
QY 595 CGAGTATGGCAAAATTAATCCATTTCTTTTAAATACGGG----AATATGGCAATTATA 650
DB 481 GCTATTGTCAAAATTAATATCCATTTCTTTTATATACGGTGAATATTCGCAATTATA 540
QY 651 GATCTGGATTTTG-ACCCTTAATGAGCGGC-ACCCAGGTGTTTGGAGTGGTGGCAT 708
DB 541 GATCTGGATTTTGAACCACTTAATGAAGCGGCAACACAGGTGTTTGGAGTGGTGGCAT 600
QY 709 TCTTCGCTGATTTGGCTGTCCCAATGTTTACATTTATTAATCTTGCAAAATGTTCTG 768
DB 601 TCTTCGCTGATTTGGCTGTCCCAATGTTTACATTTATTAATCTTGCAAAATGTTCTG 660
QY 769 TCACCTTGGATGAAATGCTGTCCAGTTTATTTTTTATATGTTTATCTTCGATGT 828
DB 661 TGCACCTTGGATGAAATGCTGTCCAGTTTATTTTTTATATGTTTATCTTCGATGT 720
QY 829 ACACAAAAATTCAGAAAAATGATCTCTGTAGATATTCGTTTATTTTGGTCATCTTAGAA 888
DB 721 ACACAAAAATTCAGAAAAATGATCTCTGTAGATATTCGTTTATTTTGGTCATCTTAGAA 780

QY 889 GTTATCAGGAATGTGTTTTAAACCAAGAGAGAACTTTTCTAAGGAATGATACATAGAAA 948
DB 781 GTTATCAGGAATGTGTTTTAAACCAAGAGAGAACTTTTCTAAGGAATGATACATAGAAA 840
QY 949 GATTTTATTTTAAATAGATTGTAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCC 1008
DB 841 GATTTTATTTTAAATAGATTGTAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCC 900
QY 1009 CAAAGTTAATGCAATGGACACATTTTATGTACAAAAACACACACACACACACACA 1068
DB 901 CAAAGTTAATGCAATGGACACATTTTATGTACAAAAACACACACACACACACACA 960
QY 1069 CACACACACACACACACGAAAAAATAAAAAAATAAAAAA 1104
DB 961 CACACACACACACACACGAAAAAATAAAAAAATAAAAAA 996
RESULT 5
ADD90210
ID ADD90210 standard; cDNA; 1510 BP.
XX AC ADD90210;
XX AC ADD90210;
DT 29-JAN-2004 (first entry)
XX Novel human secreted protein cDNA seq id 25.
DE gene therapy; cytostatic; cancer; human; gene; ss.
KW Homo sapiens.
OS
XX
XX US2003199683-A1.
XX 23-OCT-2003.
XX 30-MAR-2001; 2001US-00820649.
XX 30-JUL-1997; 97US-00542039P.
PR 30-JUL-1997; 97US-0054211P.
PR 30-JUL-1997; 97US-0054212P.
PR 30-JUL-1997; 97US-0054213P.
PR 30-JUL-1997; 97US-0054214P.
PR 30-JUL-1997; 97US-0054215P.
PR 30-JUL-1997; 97US-0054217P.
PR 30-JUL-1997; 97US-0054218P.
PR 30-JUL-1997; 97US-0054234P.
PR 30-JUL-1997; 97US-0054236P.
PR 18-AUG-1997; 97US-0055968P.
PR 18-AUG-1997; 97US-0055969P.
PR 18-AUG-1997; 97US-0055972P.
PR 19-AUG-1997; 97US-0056534P.
PR 19-AUG-1997; 97US-0056543P.
PR 19-AUG-1997; 97US-0056554P.
PR 19-AUG-1997; 97US-0056561P.
PR 19-AUG-1997; 97US-0056727P.
PR 19-AUG-1997; 97US-0056729P.
PR 19-AUG-1997; 97US-0056730P.
PR 29-JUL-1998; 98WO-US015949.
PR 26-JAN-1999; 99US-00236557.
PR 21-SEP-2000; 2000US-00666987.
XX (RUBE/) RUBEN S M.
PA (FENG/) FENG P.
PA (LAPL/) LAFLEUR D W.
PA (MOOR/) MOORE P A.
PA (SHIY/) SHI Y.
PA (KYAW/) KYAW H.
PA (LIY/) LI Y.
PA (ZENG/) ZENG Z.
PA (CART/) CARTER K C.
PA (ENDR/) ENDRESS G A.
PA (WEIY/) WEI Y.

PA (FANP/) FAN P.
PA (ROSE/) ROSEN C A.
XX
PI Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX
DR WPI; 2003-852813/79.
DR P-PSDB; ADD90301.
XX
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g., cancer.
XX
PS Claim 1; SEQ ID NO 25; 213pp; English.
XX
CC The invention describes novel isolated human nucleic acids. The nucleic
CC acid is useful for preparing a medicament for preventing, treating or
CC ameliorating a medical condition e.g., cancer, and in gene therapy. This
CC sequence encodes a novel human secreted protein of the invention.
XX
SQ Sequence 1510 BP; 473 A; 285 C; 299 G; 450 T; 0 U; 3 Other;

Query Match 63.4%; Score 699.8; DB 10; Length 1510;
Best Local Similarity 89.7%; Pred. No. 2.7e-135;
Matches 893; Conservative 1; Mismatches 69; Indels 33; Gaps 12;

QY 141 GCGGAGCTGGCCGAGATCTTCTCGCGCTGCGGAGCAGCAACCACTTCTTCGGCGCC 200
DB 2 GCGGAGCTGGCCGAGATCTTCTCGCGCTGCGGAGCAGCAACCACTTCTTCGGCGCC 61

QY 201 GCGGAGCAAGCGCGCCCAAGCTGGGCGAGATCGCGGAGCAAGCGGTGTATT 260
DB 62 GCGGAGCAAGCGCGCG-CCAGCTGGGCGAGATCGCGGAGCAAGCGGTGTATT 120

QY 261 GAGATGATAGGATTGATGACGTGCTGAAATATGACCGACAGGCACT-CTGTGTC 319
DB 121 GAGATGATAGGATTGATGACGTGCTGAAATATGACCGACAGGCACTCTCTGTGTC 180

QY 320 TAACT-CCCCAAGACAATGAGTTAAGGAGAGAGATAGGAACGGCGGTAACTGATTGG 378
DB 181 TAACTCCCCAAGACAATGAGTTAAGGAGAGAGATAGGAACGGCGGTAACTGATTGG 240

QY 379 CAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTAAGTTG-TACCCACGATGA 437
DB 241 CAAAAAGCATGAAAAGAGAAAGCACTTTGAAATTTATTAAGTTG-TACCCACGATGA 300

QY 438 AATCAACACCTGTATCTGTGTAT-ATGCCGGGAGAGATAGGCGA-AGGAGGAGAGA 495
DB 301 AATCAACACCTGTATCTGTGTATCAGCGCGGAGAGACAGATGAGGAGAGGAGGAGG 360

QY 496 GAGAGAAGAAAGCTTTGGGC-CCTCTACAAATAAATAAATAAATAAATAAATAAATAA 554
DB 361 AGGAGGAGAGGCTCTGGGCTCTCTGCAAAATAAATAAATAAATAAATAAATAAATTT 420

QY 555 AAAATCCCTATA-----TCCCATATAAGATAAAGAGTCTCAGT----- 594
DB 421 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 480

QY 595 GCAGTATTGGCRAAATTAATCCATTTCTTTTATACGGG-----AATATTGGCATTATA 650
DB 481 GCTATTGTGCAAAATTAATCCATTTCTTTTATACGGGGAATATTGGCAATTATA 540

QY 651 GATCTGGATTTTG-ACCACCTTAATGAAGCGGC-ACCCAGGTGTTTGTAGGTGTGGCAT 708
DB 541 GATCTGGATTTTGAACCACTTAATGAAGCGGCACACACAGGTGTTTGTAGGTGTGGCAT 600

QY 709 TCTTCCTGATTTGGCTGTCCCAATGTTTACATTTAATCTTTCGCAAAAATGTTCTG 768
DB 601 TCTTCCTGATTTGGCTGTTCCTCAATGTTTACATTTAATCTTTCGCAAAAATGTTCTG 660

QY 769 TGCACCTTGGATGGAATCTGTCCAGTTTATTTTTTATGTTGTTATCTTCGAGTGT 828
DB 661 TGCACCTTGGATGGAATGCTGTCCAGTTTATTTTTTATGTTGTTATCTTCGAGTGT 720

QY 829 ACACAAAATTCAGAAAATGATCTCTGTAGATATTCGTTTATTTTGGTCATCTTTAGAA 888
DB 721 ACACAAAATTCAGAAAATGATCTCTGTAGATATTCGTTTATTTTGGTCATCTTTAGAA 780

QY 889 GTTATCAGGAATGTGTTTAAAAACAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAA 948
DB 781 GTTATCAGGAATGTGTTTAAAAACAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAA 840

QY 949 GATTTTATTTAAATGAGTTGTAAGCTGTGTTTCTTTGTTGTCGAAGCTATCTGCC 1008
DB 841 GATTTTATTTAAATGAGTTGTAAGCTGTGTTTCTTTGTTGTCGAAGCTATCTGCC 900

QY 1009 CAAGTTAATGCAATGGACACATTTTATCTCGAARAAACACACACACACACACA 1068
DB 901 CAAGTTAATGCAATGGACACATTTTATGTCTGAAAAACACACACACACACACA 960

QY 1069 CACACACACACACACGAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 1104
DB 961 CACACACACACACACACACACACGAAAAAATAAATAAATAAATAAATAAATAA 996

RESULT 6
ADG90029
ID ADG90029 standard; cDNA; 1510 BP.
XX
AC ADG90029;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human cDNA from secreted protein gene 15.
XX
KW Secreted protein; gene therapy; neural disorder; immune system disorders;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; cancer; systemic lupus erythematosus;
KW rheumatoid arthritis; multiple sclerosis; thyroiditis; anaemia;
KW Grave's disease; diabetes; hepatitis; asthma; allergy; nephritis;
KW Parkinson's disease; Alzheimer's disease; atherosclerosis;
KW myocardial infarction; AIDS; infection; human; ss; gene.
XX
OS Homo sapiens.
XX
PN US2003166541-A1.
XX
PD 04-SEP-2003.
XX
PF 04-JUN-2002; 2002US-00160162.
XX
PR 30-JUL-1997; 97US-0054209P.
PR 30-JUL-1997; 97US-0054211P.
PR 30-JUL-1997; 97US-0054212P.
PR 30-JUL-1997; 97US-0054213P.
PR 30-JUL-1997; 97US-0054214P.
PR 30-JUL-1997; 97US-0054215P.
PR 30-JUL-1997; 97US-0054217P.
PR 30-JUL-1997; 97US-0054218P.
PR 30-JUL-1997; 97US-0054234P.
PR 30-JUL-1997; 97US-0054236P.
PR 18-AUG-1997; 97US-0055968P.
PR 18-AUG-1997; 97US-0055969P.
PR 18-AUG-1997; 97US-0055972P.
PR 19-AUG-1997; 97US-0056534P.
PR 19-AUG-1997; 97US-0056543P.
PR 19-AUG-1997; 97US-0056561P.
PR 19-AUG-1997; 97US-0056727P.
PR 19-AUG-1997; 97US-0056729P.
PR 19-AUG-1997; 97US-0056730P.
PR 29-JUL-1998; 98WO-US015949.
PR 26-JAN-1999; 99US-00236557.
PR 05-JUN-2001; 2001US-0295558P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX	Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI	Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX	WPI; 2003-874923/81.
DR	P-PSDB; ADG90120.
XX	Nucleic acids encoding 83 secreted polypeptides, useful for preventing,
PT	diagnosing and treating disorders related to their aberrant expression
PT	and activity.
XX	Claim 1; SEQ ID NO 25; 308pp; English.
XX	The invention relates to an isolated nucleic acid molecule encoding a
CC	secreted protein that is at least 95% identical to a polynucleotide
CC	fragment of any of the nucleotide sequences listed in table 1A of the
CC	specification, which is hybridisable to the nucleotide sequences, a
CC	polynucleotide encoding a polypeptide (or a polypeptide fragment, domain
CC	or epitope of any of the amino acid sequences) listed in table 1A of the
CC	specification, a polynucleotide which is an (allelic) variant of the
CC	nucleotide sequences listed in the specification, a polynucleotide which
CC	encodes a species homologue of the above amino acid sequences, a
CC	polynucleotide capable of hybridising under stringent conditions to any
CC	of the above polynucleotides, where the polynucleotide does not hybridise
CC	under stringent conditions to a nucleic acid molecule having a nucleotide
CC	sequence of only A or T residues. Also included are a recombinant vector
CC	comprising the above nucleic acid molecule, making a recombinant host
CC	cell comprising the above nucleic acid molecule, an isolated polypeptide
CC	comprising a sequence that is at least 95% identical to the polypeptide
CC	(or its fragment, domain, epitope, secreted form, (allelic) variant or
CC	homologue) encoded by the above nucleic acid molecule, an isolated
CC	antibody that binds specifically to the above polypeptide, a recombinant
CC	host cell produced by the above method and that expresses the above
CC	polypeptide, making an isolated polypeptide, preventing, treating or
CC	ameliorating a medical condition, diagnosing a pathological condition or
CC	a susceptibility to a pathological condition in a subject, identifying a
CC	binding partner to the above polypeptide, the gene corresponding to the
CC	cDNA sequence given in the specification, and identifying an activity in
CC	a biological assay. The nucleic acid molecule and polypeptide are useful
CC	in diagnosing, preventing, prognosing or treating diseases or disorders
CC	associated with aberrant expression and/or activity of the above
CC	polypeptide, such as neural disorders, immune system disorders, muscular
CC	disorders, reproductive disorders, gastrointestinal disorders, pulmonary
CC	disorders, cardiovascular disorders, renal disorders, proliferative
CC	disorders and/or cancers. In particular, these diseases are systemic
CC	lupus erythematosus, rheumatoid arthritis, multiple sclerosis, asthma,
CC	thyroiditis, anaemia, Grave's disease, diabetes, hepatitis, Alzheimer's
CC	disease, nephritis, Parkinson's disease, Alzheimer's disease,
CC	atherosclerosis, myocardial infarction, AIDS and infections. The methods
CC	may be used for identifying agonists and antagonists of the
CC	polynucleotide and polypeptide. The present sequence is a cDNA from one
CC	of the 83 disclosed secreted protein genes.
XX	Sequence 1510 BP; 473 A; 285 C; 299 G; 450 T; 0 U; 3 Other;
SQ	Query Match 63.4%; Score 699.8; DB 10; Length 1510;
	Best Local Similarity 89.7%; Pred. No. 2.7e-135;
	Matches 893; Conservative 1; Mismatches 69; Indels 33; Gaps 12
QY	141 GGCGACGTGGGCCAGATCTTCTCCTCGCGCTGCAGGACACCACAACCTTTCTTGCGGCC 200
Db	2 GGCGACGTGGGCCAGATCTTCTCCTCGCGCTGCAGGNCAACCACAACCTTTCTTGCGGCC 61
QY	201 GGGCAGAACAGCGCGCCCCCAAGCTGGCCAGATCGCGCGGAGCAAGCGGGTGTATT 260
Db	62 GGGCAGAACAGCGCGCG - CCNAGCTGGCGGAGATCGCGCGGAGCAAGCGGGTGTATT 120
QY	261 GAAGATGATAGGATTGATGACTGCTGAAAAATATGACCGACAGCGACCT - CTGGTGTTC 319
Db	121 GAAGATGATAGGATTGATGACTGCTGAAAAATATGACCGACAGCGACCTCTCGTGTTC 180
QY	320 TAACAT - CCCCAAAGACAATGAGTTAAAGGAGAGAATAGGAACGGCGGTAAACAGTTATTGG 378

```
XX 26-NOV-2003; 2003WO-US038193.
XX PF
XX PR
XX 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX Example 2; SEQ ID NO 7058; 210pp; English.
XX The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytostatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 2058 BP; 488 A; 468 C; 405 G; 695 T; 0 U; 2 Other;
Query Match 58.2%; Score 642; DB 12; Length 2058;
Best Local Similarity 87.3%; Pred. No. 2.9e-123;
Matches 838; Conservative 0; Mismatches 90; Indels 32; Gaps 11;
QY 177 GACACCAACAACTTCTTCGGCGCGGCGGAGAACAGCGCCGCCCAAGCTGGCCAGATC 236
DB 1873 GAGACAGAACTCGATCGCTGCGCGCGCTGCGCGGAGGACCAAGCTGGCCAGATC 1814
QY 237 GCGCGGAGCAAGCGGCTTGTATTGAAGATGATGAGATTGATGCTGCTGAAAAATATG 296
DB 1813 GCGCGGAGCAAGCGGCTTGTATTGAAGATGATGAGATTGATGCTGCTGAAAAATATG 1754
QY 297 ACCGACAGGCACCT-CTGCTGCTTAAC-CCCCAAGCAATGATGATTAGGCGAGAT 354
DB 1753 ACCGACAGGCACCTCTGCTGCTTAACCTCCCAAGCAATGATGATTAGGCGAGAT 1694
QY 355 AGGAACGGCGGTAACTGTTATTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTA 414
DB 1693 AAGAAGCGGGTAACAGTTATTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTA 1634
QY 415 TTACTAGCTTG-TACCAAGATGAATCAACACCTGTATCTGGTAT-ATGCCCGGAGAC 472
DB 1633 TTACTAGCTTGTTACCAAGATGAATCAACACCTGTATCTGGTATCAGCGCGGAGAC 1574
QY 473 AGATTAGGCGA-AGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 530
DB 1573 AGATTAGGCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1514
QY 531 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 574
DB 1513 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1454
QY 575 AGAATAAAGAGTCTCACT-CCAGTATTTGGCAAAATTAATTAATTAATTAATTAAT 630
DB 1453 AATAAATAAAGAGTCTCACTTTCAGCTATTTTCAAAATTAATTAATTAATTAATTA 1394
QY 631 ACCGGAATATT---GGCAATTATAGATCTGGAATTTG-ACCACCTTAATGAAGCGGC-ACC 684
DB 1393 TAGCGTGAATATTGGCAATTAATGATCTGGATTTTGAACCACTTAATGAAGCGGCAACA 1334
```

```
QY 685 CCAGGTGTTTGGAGTGTGGCATCTCTGCTGATTTGGCTGTTCCCAATGTTTACATTA 744
DB 1333 CCAGGTGTTTGGAGTGTGGCATCTCTGCTGATTTGGCTGTTCCCAATGTTTACATTA 1274
QY 745 TTTAATCTTGCAGAAATGTTCTGTCGACTTGGATGTGAAATGCTGTCCAGTTTATTTT 804
DB 1273 TTTAATCTTGCAGAAATGTTCTGTCGACTTGGATGTGAAATGCTGTCCAGTTTATTTT 1214
QY 805 TTTTATGTTTATCTCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATTTCT 864
DB 1213 TTTTATGTTTATCTCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATTTCT 1154
QY 865 GTTTTATTTTGGTCATCTTTAGAGTATATCAGGAATGTTTAAACCAAGAGAGAACTT 924
DB 1153 GTTTTATTTTGGTCATCTTTAGAGTATATCAGGAATGTTTAAACCAAGAGAGAACTT 1094
QY 925 TTCTAAGGAATGATACATAGAAAAAGATTTTATTTTAAATGATTTGTAAGCTTGTGTTT 984
DB 1093 TTCTAAGGAATGATACATAGAAAAAGATTTTATTTTAAATGATTTGTAAGCTTGTGTTT 1034
QY 985 CTTTGTCTGCAAGCTATCTGCCCAGATTAAATGCAAAATGGACACATTTTATGTCTAGA 1044
DB 1033 CTTTGTCTGCAAGCTATCTGCCCAGATTAAATGCAAAATGGACACATTTTATGTCTAGA 974
QY 1045 AAAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 1104
DB 973 AAAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 914
RESULT 8
ACN90200
ID ACN90200 standard; DNA; 1850 BP.
XX AC ACN90200;
XX 02-DEC-2004 (first entry)
XX Breast cancer related marker, seq id 11350.
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX Homo sapiens.
XX US2003099974-A1.
XX 29-MAY-2003.
XX 18-JUL-2002; 2002US-00198846.
XX 18-JUL-2001; 2001US-0306220P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2003-787014/74.
XX Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.
XX Disclosure; SEQ ID NO 11350; 36pp; English.
XX The invention relates to an isolated polypeptide (I) associated with a
XX breast cancer which is encoded by a nucleic acid molecule comprising a
XX nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX the polypeptide of the invention. The activity of the polypeptide of the
XX invention may be described as cytostatic. The antibody is useful for
XX detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX invention are useful in the detection of breast tumours. (I) is useful as
XX a marker for breast cancer and in breast cancer therapy. Sequences given
XX in records ACN78851-ACN92934 represent nucleic acid markers associated
```


Query Match

57.3%; Score 632.6; DB 12; Length 1321;

PF 11-OCT-2001; 2001WO-US027760.
XX
PR 12-OCT-2000; 2000US-00687527.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-426278/45.
XX N-PSDB; ABP43731.
XX
PT New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
PS Claim 1; SEQ ID # 188; 357pp + Sequence Listing; English.
XX
CC The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnerability, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records ABQ60788-
CC ABQ1233 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 873 BP; 304 A; 144 C; 174 G; 251 T; 0 U; 0 Other;

Query Match 47.1%; Score 519.8; DB 6; Length 873;
Best Local Similarity 88.2%; Pred. No. 5.1e-98;
Matches 721; Conservative 0; Mismatches 62; Indels 34; Gaps 13;

QY 299 GCACGAGGACCT-CTGGTGTCTAACT-CCCCAAGACAACTAGTTAAGGGAGAGAAATAG 356
DB 53 GCACGAGGACCTCTCTGGTGTCTAACTCCCCAAGACAACTAGTTAAGGGAGAGAAATAA 112
QY 357 GAACGCGGTAAACAGTTATTGGCAAAAGCATGAAAGAGAACGACCTTTGAAATTTATT 416
DB 113 GAACGCGGTAAACAGTTATTGGCAAAAGCATGAAAGAGAACGACCTTTGAAATTTATT 172
QY 417 ACTAGCTTG-TACCCACGATGAAATCAACACCTGTATCTGTAT-ATGCCCGGAGACAG 474
DB 173 ACTAGCTTGCTACCCACGATGAAATCAACACCTGTATCTGTATCAGGCGGAGACAG 232
QY 475 ATTAGCGGA-AGGAGGAGAGAGAGAGAGAGAGAGGCTTGGGC-CTCTACAAATAATA 532
DB 233 ATGAGCGGAGAGGAGGAGAGAGAGAGAGAGGCTCTGGGCTCTCTGCAAAATAATAA 292
QY 533 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 576
DB 293 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 352
QY 577 GAATAAAGAGCTCAGT-----GCAGTATGGCAAAATTAATCCATTTCTTTTAAATAC 632
DB 353 AAAAAAAGAGCTCAGTTGCGAGCTATTGTCAAAATAAATATCCATTTCTTTTAAATATA 412
QY 633 GGG-----ATATGTGCATATATAGATCTGGATTTTG-ACCACCTAATGAGCGGC-ACCCC 686
DB 413 CGGTGAATATTCGCAATATATAGATCTGGATTTTGAAACCACTTAATGAAGCGGCACACC 472
QY 687 AGGTGTTTTAGGTGTTGCATTTCTCGCTGATTTTGGCTGTTTCCCAATGTTTACATTTATT 746
DB 473 AGGTGTTTTAGGTGTTGCATTTCTCGCTGATTTTGGCTGTTTCCCAATGTTTACATTTATT 532
QY 747 TAATCTGCAAAATGGTTCTGTGCACCTTGGATGTGAAATGCTGCCAGTTTATTTTTTT 806

DB 533 TAATCTGCAAAATGGTTCTGTGCACCTGGATGTGAAATGCTGCCAGTTTATTTTTT 592
QY 807 TTATGTTCTTATCTTGGATGTACAAAAAATTCAGAAAAATGATCTCTGTAGATATTTCTGT 866
DB 593 TTATGTTCTTATCTTGGATGTACAAAAAATTCAGAAAAATGATCTCTGTAGATATTTCTGT 652
QY 867 TTTATTTGGTCACTCTTTAGAGTTATCAGGAATGTGTTTAAACAAGAGAGAACTTTT 926
DB 653 TTTATTTGGTCACTCTTTAGAGTTATCAGGAATGTGTTTAAACAAGAGAGAACTTTT 712
QY 927 CTAAGGAATG-ATACATAG-AAAAAGATTTTATTTTAAATGAGTTGTTAAAGCTTGTGTTT 984
DB 713 CTAAGGAATGATACATAGCAAAAGATTTTATTTTAAATGAGTTGTTAAAGCTTGTGTTT 772
QY 985 CTTTGTCTGCAAGCTATCTGCCCAAGTTAATGCAAAATGACACATTTTTTTATGTCAGA 1044
DB 773 CTTTGTCTGCAAGCTATCTGCCCAAGTTAATGCAAAATGACACATTTTTTTATGTCAGA 832
QY 1045 AAAACACACACACACACACACACACACACACACACACAC 1081
DB 833 AAAACACACACACACACACACACACACACACACACACAC 869

RESULT 14
ABX97604
ID ABX97604 standard; cDNA; 553 BP.
XX
AC ABX97604;
XX
DT 16-MAY-2003 (first entry)
XX
DE cDNA encoding novel human protein NOV9a.
XX
KW Human; NOV; adrenoleukodystrophy; congenital adrenal hyperplasia;
KW haemophilia; hypercoagulation; autoimmune disease; allergy;
KW immunodeficiency; transplantation; Von Hippel-Lindau syndrome;
KW Alzheimer's disease; stroke; tuberculous sclerosis; hypercalcaemia;
KW Parkinson's disease; Huntington's disease; cancer; fertility; diabetes;
KW adult respiratory distress syndrome; infection; tissue typing;
XX
XX forensic identification; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200290500-A2.
XX
PD 14-NOV-2002.
XX
PF 02-MAY-2002; 2002WO-US014256.
XX
PR 03-MAY-2001; 2001US-0288395P.
PR 07-MAY-2001; 2001US-0289087P.
PR 08-MAY-2001; 2001US-0289619P.
PR 09-MAY-2001; 2001US-0289817P.
PR 09-MAY-2001; 2001US-0289818P.
PR 11-MAY-2001; 2001US-0290194P.
PR 14-MAY-2001; 2001US-0290753P.
PR 15-MAY-2001; 2001US-0291189P.
PR 21-MAY-2001; 2001US-0292374P.
PR 23-MAY-2001; 2001US-0293107P.
PR 25-MAY-2001; 2001US-0293747P.
PR 29-MAY-2001; 2001US-0294110P.
PR 30-MAY-2001; 2001US-0294434P.
PR 10-SEP-2001; 2001US-0318346P.
PR 17-SEP-2001; 2001US-0322646P.
PR 01-MAY-2002; 2002US-00136728.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Spytek KA, Li L, Edinger SR, Stone DJ, Guo X, Anderson DW;
PI Patturajan M, Gerlach VL, Taupier RJ, Pena CE, Padigar M;
PI Kekuda R, Gorman L, Zehrhusen BD, Smithson G, Macdougall JR;
PI Mezes PS, Peyman JA, Zhong M;

XX WPI; 2003-103511/09.
DR P-PSDB; ABU65229.
XX
XX New NOVX polypeptides and polynucleotides useful for treating or
PT preventing e.g. congenital adrenal hyperplasia, hemophilia,
PT hypercoagulation, autoimmune disease, allergies, immunodeficiencies,
PT transplantation.
XX
XX Claim 1; Page 106-107; 300pp; English.
XX
XX The invention describes an isolated polypeptide, NOVX, comprising a
CC sequence or a mature form of one of 21 51-1543 residue amino acid
CC sequences (p1-p21), given in the specification. The NOVX polypeptides,
CC polynucleotides and antibodies are useful in the manufacture of a
CC medicament for treating or preventing e.g. adrenoleukodystrophy,
CC congenital adrenal hyperplasia, haemophilia, hypercoagulation, autoimmune
CC disease, allergies, immunodeficiencies, transplantation, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cancer,
CC fertility, diabetes, adult respiratory distress syndrome, viral
CC bacterial and parasitic infections. The nucleic acid sequences may be
CC used in chromosome mapping, identifying individual from minute biological
CC samples (tissue typing), and in forensic identification of a biological
CC sample. This sequence encodes a novel human protein (NOV)
XX
XX Sequence 553 BP; 169 A; 133 C; 170 G; 81 T; 0 U; 0 Other;
SQ
Query Match 40.1%; Score 443.2; DB 8; Length 553;
Best Local Similarity 94.0%; Pred. No. 3.4e-82;
Matches 515; Conservative 0; Mismatches 28; Indels 5; Gaps 5;
QY 16 CGCTGACCTCTGTCGGCGGGCGGAGCGCGGGCGGAGGCGCGGGCGGAGCGCC 75
Db 1 CGCTGACCTCTGTCGGCGGGCGGAGCGCGGGCGGAGGCGCGGGCGGAGCGCC 60
QY 76 CGAGCGGACCATGTGGAGGTCTGCCCTAOCGCGACGAGAGCTGAGCCCTACGGCG 135
Db 61 CGGACGCGACATGTGCGAGGTCTGCCCTAOCGCGACGAGAGCTGAGCCCTACGGCG 120
QY 136 ACGGCGGCGAGCTGGCGCCAGATCTCTCTGCGCTGCGCGGCGGAGCACACCACTCTTCG 195
Db 121 ACGGCGGCGAGCTGGCGCCAGATCTCTCTGCGCTGCGCGGCGGAGCACACCACTCTTCG 180
QY 196 GCGCCGGCGAGAACAGCGCGCCGCCAAGCTGGCGCAGATCGCGCGGAGCAAGCGGTTG 255
Db 181 GCSCCGGCGAGAACAGCGCGCCGCCAAGCTGGCGCAGATCGCGCGGAGCAAGCGGTTG 240
QY 256 TTATTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314
Db 241 TTATTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 315 GTGTCTAACT-CCCCAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373
Db 301 GTGTCTAACTCCCCAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 374 ATTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTG-TACCCAC 432
Db 361 ATTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGCTACCCAC 420
QY 433 GATGAATCAGCAACTGATCTGGTAT-ATGCCCGGAGACAGATGAGCGA-AGAGGA 490
Db 421 GATGAATCAGCAACTGATCTGGTATCAGGCGCGGAGACAGATGAGCGGAGAGGA 480
QY 491 AGAGAGAGAGAGAGAGAGGCTTGGGCCCTCTCAAAATAAAATAAAATAAAATAAAATAAA 550
Db 481 GGAGGAGAGAGAGAGAGGCTTGGGGCTCTCTGCAAAAAAAAAAAAAAAAAAAAAA 540
QY 551 TAATAAAA 558
Db 541 AAAAAAAA 548

RESULT 15
AAC06588
ID AAC06588 standard; cDNA; 524 BP.
XX
AC AAC06588;
XX
DT 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 10663.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 10663; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
XX Sequence 524 BP; 159 A; 86 C; 93 G; 179 T; 0 U; 7 Other;
SQ
Query Match 38.7%; Score 427.2; DB 3; Length 524;
Best Local Similarity 97.8%; Pred. No. 7e-79;
Matches 451; Conservative 2; Mismatches 6; Indels 2; Gaps 2;
QY 645 ATTATAGATCTGGATTTTG-ACCATTAAATGAAGCGC-ACCCAGGTGTTTTCAGGTGT 702
Db 24 ATTATAGATCTGGATTTTGACCACTTAATGAAGCGCACACCGGTGTTTCAGGTGT 83
QY 703 TGGCATTTCTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTTGCAAAATG 762
Db 84 TGGCATTTCTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTTGCAAAATG 143
QY 763 GTTCGTGCACTTCGATGTGAATGCTGTCAGTTTTATTTTTTTTATGTTGTTATCCTT 822
Db 144 GTTCGTGCACTTCGATGTGAATGCTGTCAGTTTTATTTTTTTTATGTTGTTATCCTT 203
QY 823 GGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTTCTGTTTATTTTGTGTCATCT 882
Db 204 GGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTTCTGTTTATTTTGTGTCATCT 263
QY 883 TTAGAAGTTATCAGGAATGTGTTTAAACAAGAGAGAACTTTTCTAAGGATGATACAT 942


```
Db 264 TTAGAAGTTATCAGGAATGTGTTTAAAAACAGAGAGAACTTTTCTAAGGAATGATACAT 323
Qy 943 AGAAAAGATTTTATTTTAAATGAGTTGTAAAGCTTGTGTTTCTTTTGTGCTGCAAGCTA 1002
Db 324 AGAAAAGATTTTATTTTAAATGAGTTGTAAAGCTTGTGTTTCTTTTGTGCTGCAAGCTA 383
Qy 1003 TCTGCCCAAGTTAATGCAAAATGGACACATTTTTTATGTCAGAAAAACACACACACACA 1062
Db 384 TCTGCCCAAGTTAATGCAAAATGGACACATTTTTTATGTCAGAAAAACACACACACACA 443
Qy 1063 CACACACACACACACACACGAAAAA 1103
Db 444 CACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 484
```

Search completed: March 30, 2005, 00:47:34
Job time : 643 secs

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description	
1	803.8	72.8	1876	4	US-09-566-921-70	Sequence 70, Appl	
2	790	71.6	1581	3	US-09-313-300-1	Sequence 1, Appli	
3	427.2	38.7	524	4	US-09-513-999C-10663	Sequence 10663, A	
C 4	270	24.5	417	4	US-08-702-705-475	Sequence 475, App	
C 5	270	24.5	417	4	US-09-736-457-475	Sequence 475, App	
C 6	270	24.5	417	4	US-09-614-124B-475	Sequence 475, App	
C 7	270	24.5	417	4	US-09-671-325-475	Sequence 475, App	
C 8	270	24.5	417	4	US-09-589-184-475	Sequence 475, App	
C 9	270	24.5	417	4	US-09-658-824-475	Sequence 475, App	
10	165	14.9	336	4	US-09-854-133-560	Sequence 560, App	
11	155.8	14.1	183	4	US-09-513-999C-10536	Sequence 10536, A	
12	118.6	10.7	1045	3	US-09-313-300-9	Sequence 9, Appli	
13	96.4	8.7	315	4	US-09-621-976-15158	Sequence 15158, A	
14	94	8.5	94	4	US-09-513-999C-29981	Sequence 29981, A	
15	77.6	7.0	90	4	US-09-513-999C-24816	Sequence 24816, A	
C 16	67.2	6.1	7218	1	US-08-232-463-14	Sequence 14, Appl	
C 17	60.2	5.5	111509	4	US-09-949-016-17379	Sequence 17379, A	
C 18	58.2	5.3	601	4	US-09-949-016-48404	Sequence 48404, A	
19	58	5.3	92139	4	US-09-918-686-1	Sequence 1, Appli	
C 20	57.6	5.2	601	4	US-09-949-016-129237	Sequence 129237, A	
C 21	57.6	5.2	98864	4	US-09-949-016-15403	Sequence 15403, A	
C 22	57.6	5.2	114842	4	US-09-949-016-14993	Sequence 14993, A	
C 23	56.8	5.1	360470	4	US-09-949-016-13173	Sequence 13173, A	
C 24	56.4	5.1	601	4	US-09-949-016-48403	Sequence 48403, A	
C 25	56	5.1	102304	4	US-09-949-016-12589	Sequence 12589, A	
C 26	55.6	5.0	28139	4	US-09-949-016-17168	Sequence 17168, A	
C 27	55.6	5.0	28129	4	US-09-949-016-17169	Sequence 17169, A	

[illegible]

RESULT 2

```

US-09-313-300-1
; Sequence 1, Application US/09313300
; Patent No. 6222027
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew, R.
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Tang, Tom, Y.
; APPLICANT: Baughn, Mariah, R.
; APPLICANT: Azimzai, Valda
; TITLE OF INVENTION: MOLECULES EXPRESSED IN HIPPOCAMPUS
; FILE REFERENCE: PB-0012 US
; CURRENT APPLICATION NUMBER: US/09/313,300
; CURRENT FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 15

```

[illegible]

RESULT 3

```

RESULT 3
; US-09-513-999C-10663
; Sequence 10663, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

```

```

, Patent NO. 6783961
, FILE REFERENCE: 59.US2.REG
, CURRENT APPLICATION NUMBER: US/09/513,999C
, CURRENT FILING DATE: 2000-02-24
, PRIOR APPLICATION NUMBER: US 60/122,487
, PRIOR FILING DATE: 1999-02-26
, NUMBER OF SEQ ID NOS: 36681
, SOFTWARE: Patent.pm
, SEQ ID NO 10663
, LENGTH: 524

```

```

; LIFE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 465
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 471
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:

```

Query Match 38.7%; Score 427.2; DB 4; Length 524;
Best Local Similarity 97.8%; Pred. No. 2.4e-91;
Matches 451; Conservative 2; Mismatches 6; Indels 2

645	Qy	ATTATAGATCTGATTTTG-ACCACTTAATGAAGCGC-ACCCACGGTGTTTTGAGGTGT	702
24	Db	ATTATAGATCTGATTTTGAACCACTTAATGAAGCGGCAACACAGGTGTTTTGAGGTGT	83
703	Qy	TGSCATTTCTTGGCTGATTTGGCTGTTCCTCAATGTTTACATTAATTAATCTTGCAAAAATG	762
84	Db	TGSCATTTCTTGGCTGATTTGGCTGTTCCTCAATGTTTACATTAATTAATCTTGCAAAAATG	143
763	Qy	GTTCTGTGCACCTTGGATGTGAAATGCTGTCCAGTGTTTTATTTTTTATGTGTATATCTT	822
144	Db	GTTCTGTGCACCTTGGATGTGAAATGCTGTCCAGTGTTTTATTTTTTATGTGTATATCTT	203
823	Qy	GGATGTACAAAAAATTCAGAAAAATGATCTCTGTAGATATTCCTGTTTTATTTTGGTCACT	882
204	Db	GGATGTACAAAAAATTCAGAAAAATGATCTCTGTAGATATTCCTGTTTTATTTTGGTCACT	263
883	Qy	TTAGAAGTTATCAGGAATGTGTTTAAAAACAAGAAGAGAACTTTTCTAAGAAATGATACAT	942
264	Db	TTAGAAGTTATCAGGAATGTGTTTAAAAACAAGAAGAGAACTTTTCTAAGAAATGATACAT	323
943	Qy	AGAAAAGATTTTATTTTAAAAATGAGTTGTAAAGCTTGTGTTCTCTTGTCTGCCAGCTA	1002
324	Db	AGAAAAGATTTTATTTTAAAAATGAGTTGTAAAGCTTGTGTTCTCTTGTCTGCCAGCTA	383
1003	Qy	TCGTGCCCAAGTTAAATGCAAAATGGACACATTTTTTATGTTCAGAAAAACACACACACACA	1062
384	Db	TCGTGCCCAAGTTAAATGCAAAATGGACACATTTTTTATGTTCAGAAAAACACACACACACA	443
1063	Qy	CACACACACACACACACACACGAAAAAATAAAAAAATAAAAAA 1103	
444	Db	CACACACACACACACACACACGCGAANAACAGAAAAAATAAAAAA 484	

RESULT 4

```

US-09-702-705-475/C
; Sequence 475, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-702-705-475

```

Query Match	24.5%	Score 270;	DB 4;	Length 417;
Best Local Similarity	83.4%	Pred.No. 3.2e-54;		
Matches 342;	Conservative	0;	Mismatches 62;	Indels 6; Gaps 3
521	ACAAATAAAATAAAAAAATTTTAAATAATAAAAAATCCCTATATCCCATATAAGAAT	580		
410	ATAAATAAAATTTTAAATAATAAAAAATTCACATATATNCNCATATAAGAAATAAAAAAG	351		

Query Match	38.7%	Score 427.2;	DB 4;	Length 524;
Best Local Similarity	97.8%	Pred. No. 2.4e-91;		
Matches 451: Conservative	2;	Mismatches 6;	Indels 2;	Gaps 2;


```
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C11
; CURRENT APPLICATION NUMBER: US/09/658,824
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 1788
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-658-824-475

Query Match      24.5%; Score 270; DB 4; Length 417;
Best Local Similarity 83.4%; Pred. No. 3.2e-54;
Matches 342; Conservative 0; Mismatches 62; Indels 6; Gaps 3;

QY 521 ACAATAAATAAAAAAATTTAAATAATAAAAAATCCCTATATCCCATATAAGAAAT 580
Db 410 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAG 351
QY 581 AAAAGAGTCTCAGTCAGTATGGCAAAATTAATCCATCTCTTTTAAACGGG----A 636
Db 350 AGTCTCAGTTGCGAGCTATTGTCAAAATTAATATCCATTTCTTTTATATACGGTGAAT 291
QY 637 ATATTGGCATTATAGATCTGGATTTTG-ACCACCTAATGAAGCGGCACC-CCAGGTGTTT 694
Db 290 ATTGCGCAATATAGATCTGGAATTTTGAACCACTTAATGAAGCGGCACACCAAGTGT 231
QY 695 TGAGGTGTTGGCAATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTTG 754
Db 230 TGAGGTGTTGGCAATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTTG 171
QY 755 CAAAAATGGTCTGTCGACCTTGATGTAATGTAATGTAATGTAATGTAATGTAATGTA 814
Db 170 CAAAAATGGTCTGTCGACCTTGATGTAATGTAATGTAATGTAATGTAATGTAATGTA 111
QY 815 TTATCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATCTGTTTATTTT 874
Db 110 TTATCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATCTGTTTATTTT 51
QY 875 GGTCACTTTAGAGTTATCAGGAATGTGTTTAAACAAGAGAGAACTT 924
Db 50 GGTCACTTTAGAGTTATCAGGAATGTGTTTAAACAAGAGAGAACTT 1
```

```
RESULT 10
US-09-854-133-560
; Sequence 560, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 560
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(336)
; OTHER INFORMATION: n = A,T,C or G
US-09-854-133-560

Query Match      14.9%; Score 165; DB 4; Length 336;
Best Local Similarity 76.8%; Pred. No. 2e-29;
Matches 238; Conservative 0; Mismatches 66; Indels 2; Gaps 3;

QY 524 AATAAATAAAAAAATTTAAATAATAAAAAATCCCTATATCCCATATAAGAAATAAA 583
Db 27 AATAAATAAATTTAAATAATAAATAAATTCATATATACACATATAAAGAAATAAAGAG 86
QY 584 AGAGTCTCAGTCAGTATTTGGCAAAATTAATCCATCTCTTTTAAATACGGG----AATA 639
Db 87 TCTCAGTTGCGAGCTATTGTCAAAATTAATATCCATCTCTTTTATATACGGTGAATAT 146
QY 640 TTGGCATTATAGATCTGGATTTTG-ACCACCTAATGAAGCGGCACC-CCAGGTGTTTGA 697
Db 147 GCGCAATATAGATCTGGATTTTGAAACCACTTAATGAAGCGGCACACACAGGTGTTTGA 206
QY 698 GGTGTGTCATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTTAAATCTTGCAA 757
Db 207 GGTGTGTCATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTTAAATCTTGCAA 266
QY 758 AAATGGTCTCTGCACTTGGATGTGAATGCTGCCAGTTTATTTTATGTTGTTA 817
Db 267 AAATGGTCTCTGCACTTGGATGTGAATGCTGCCAGTTTATTTTATGTTGTTA 326
QY 818 TCCTTGGATG 827
Db 327 TCCTTGGATG 336

RESULT 11
US-09-513-999C-10536
; Sequence 10536, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 10536
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-10536

Query Match      14.1%; Score 155.8; DB 4; Length 183;
Best Local Similarity 97.8%; Pred. No. 2.2e-27;
Matches 179; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 183 AACAACTTCTTCGGCGCGCGGCAGAACAAAGCGGCGCCCAAGCTGGGCCAGATCGGCGG 242
Db 1 AACAACTTCTTCAGCGCGCGGCAGAACAAAGCGGCGCCCAAGCTGGGCCAGATCGGCGG 60
QY 243 AGCAAGCGGGTGTATTGTAAGATGATGAGATTGATGATGATGATGATGATGATGATGAT 302
Db 61 AGCAAGCGGGTGTATTGTAAGATGATGAGATTGATGATGATGATGATGATGATGATGAT 120
QY 303 AAGGCACCT-CTGGTGTCTAACT-CCCCCAAGCAATGAGTTAAGGAGAGAGATAGGAC 360
Db 121 AAGGCACCTCTCTGGTGTCTAACTCCCCCAAGCAATGAGTTAAGGAGAGAGATAGGAC 180
```

QY 361 GGC 363
|||
Db 181 GGC 183

RESULT 12

US-09-313-300-9
; Sequence 9, Application US/09313300
; Patent No. 6222027
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew, R.
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Tang, Tom, Y.
; APPLICANT: Baughn, Mariah, R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: MOLECULES EXPRESSED IN HIPPOCAMPUS
; FILE REFERENCE: PB-0012 US
; CURRENT APPLICATION NUMBER: US/09/313,300
; CURRENT FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 1045
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (871)...(899)
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY:
; OTHER INFORMATION: 700122146
; PUBLICATION INFORMATION:
US-09-313-300-9

Query Match 10.7%; Score 118.6; DB 3; Length 1045;
Best Local Similarity 62.1%; Pred. No. 3.2e-18;
Matches 259; Conservative 0; Mismatches 144; Indels 14; Gaps 4;
QY 532 AAAAAAAAAAATTTAAATAATAAATCCCTATATCCCATATAGATATAAGATGCTC 591
Db 2 AATAAATAAATAAATAAATCACTATATACACATATAAAGAAAAAGTGTGAGTTGC 61
QY 592 AGTGCAGTATGGCAAAATTAATCCATTCTTTTAAATACGGGAATAT--TGGCATTAT 649
Db 62 AGCTACTTGTGAAATTAATACCTGTTCTTTTATCTATGTAATATCGTGCATATAT 121
QY 650 AGATCTGATTTTG-ACCACCTTAATGAAGCGGCACCCAGGTGTTTGGAGGTGTTGGCAT 708
Db 122 AGATCTGATTTTGAACCACTTCTGAAAGCAGCAGAGTACTCGAAGGTGCTTGTGT 181
QY 709 TCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTAATCTTGGCAAAATGTTCTG 768
Db 182 TCTCTGCTGATTTGGCTGTTTCCCAATGTTTACATTAATCTTGGCAAAATGATCTG 241
QY 769 TGCACCTGATGAAATGCTGTCAG------TTTTATTTTTTTTATGTTTATCTT 822
Db 242 TGCACCTGATGATGACATGCTGTCTAGTCCGGTTTCATCTTTTTTTTAAATGTTTAT 301
QY 823 GGATGTACAAAAAATTCAGAAAAATGATCTCTGTAGATATCTGTTTATTTTGGTCACT 882
Db 302 TTTTGGATGTACAAAAAATAATTTGGGGGAGGGGTGATCTCTGTAGATCTCTTGTAC 361
QY 883 TTAGAGTTATC-----AGGAATGTTTAAAAACAAGAAAGAACTTTCTTAAGAA 934
Db 362 TTTGAAGTTACCGGAAATGGAACGGGTCTTAAAGCAGAAAGTAATCTTTTCCAAGAA 418

RESULT 13

US-09-621-976-15158
; Sequence 15158, Application US/09621976

; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15158
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15158

Query Match 8.7%; Score 96.4; DB 4; Length 315;
Best Local Similarity 92.5%; Pred. No. 3.2e-13;
Matches 135; Conservative 0; Mismatches 6; Indels 5; Gaps 3;
QY 252 GTTGTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATGACCGACAAGCACCT 311
Db 172 GTTGTATTGAAGATGATAGGATTGATGACG---TGAATAATATGACCGACAAGCACCT 228
QY 312 -CTGGTGTCTAACT-CCCCAAAGACAATGAGTTAAGGGAGAGAAATAGGAACGGCGTAAAC 369
Db 229 CCTGGTGTCTAACTCCCCAAAGACAATGAGTTAAGGGAGAGAAATAGGAACGGCGTAAAC 288
QY 370 AGTTATTGGCAAAAGCATGAAAAGA 395
Db 289 AGTTATTGGCAAAAGCATGAAAAGA 314

RESULT 14

US-09-513-999C-29981
; Sequence 29981, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclett, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 29981
; LENGTH: 94
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-29981

Query Match 8.5%; Score 94; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 6.6e-13;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 987 TTGTGTGTCGAAGCTATCTGCCCAAGTTAATGCAAAATGGACACATTTTTTATGTCAGAAA 1046
Db 1 TTGTGTGTCGAAGCTATCTGCCCAAGTTAATGCAAAATGGACACATTTTTTATGTCAGAAA 60
QY 1047 AAC 1080
Db 61 AAC 94

RESULT 15

US-09-513-999C-24816
; Sequence 24816, Application US/09513999C

; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 24816
; LENGTH: 90
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 23
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 70
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-24816

Query Match 7.0%; Score 77.6; DB 4; Length 90;
Best Local Similarity 96.7%; Pred. No. 4.9e-09;
Matches 88; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
Qy 976 CTTGCTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAATGGACACATTTT 1035
Db 1 CTTGCTTTCTTTGTTGCTGCTGC-ARCTATCTGCCCAAGTTAATGCAATGGACATTTT 59
Qy 1036 TATGTCAGAAAACACACACACACACACACA 1066
Db 60 TATGTCAGAAAACACACACACACACACACA 90

Search completed: March 30, 2005, 03:12:17
Job time : 233 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2005, 02:08:00 ; Search time 666 Seconds
(without alignments)
9878.532 Million cell updates/sec

Title: US-10-071-645-1
Perfect score: 1104
Sequence: 1 ggcacgagcgccgctg.....aaaaaaaaaaaaaaaaaaaaa 1104

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 297965951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1104	100.0	1104	15	US-10-071-645-1
2	699.8	63.4	1510	10	US-09-820-649-25
3	699.8	63.4	1510	16	US-10-160-162-25
4	699.8	63.4	1510	19	US-10-936-773-25
c	642	58.2	2058	18	US-10-723-860-7058
5	634.6	57.5	1850	14	US-10-198-846-11350
6	634.6	57.3	1321	17	US-10-172-118-1796
7	632.6	57.3	1321	17	US-10-342-887-1796
8	632.6	57.3	1321	17	US-10-342-887-1796
9	607.2	55.0	3189	15	US-10-071-645-1
10	443.2	40.1	553	17	US-10-136-728-21
11	405.4	36.7	417	11	US-09-978-360A-155

12	380.4	34.5	584	13	US-10-027-632-231746	Sequence 231746,
13	380.4	34.5	584	17	US-10-027-632-231746	Sequence 231746,
14	343.8	31.1	366	17	US-10-242-535A-49104	Sequence 49104, A
15	343.8	31.1	366	17	US-10-085-783A-49104	Sequence 49104, A
16	295.8	26.8	360	17	US-10-242-535A-44577	Sequence 44577, A
17	295.8	26.8	360	17	US-10-085-783A-44577	Sequence 44577, A
18	282.2	25.6	411	9	US-09-964-824A-322	Sequence 322, App
19	282.2	25.6	411	9	US-09-954-456-313	Sequence 313, App
20	282.2	25.6	411	9	US-09-954-456-2149	Sequence 2149, App
21	282.2	25.6	411	10	US-09-873-367C-677	Sequence 677, App
22	272	24.6	620	14	US-10-198-846-8710	Sequence 8710, App
23	270	24.5	417	9	US-09-736-457-475	Sequence 475, App
24	270	24.5	417	9	US-09-902-941-475	Sequence 475, App
25	270	24.5	417	9	US-09-849-626-475	Sequence 475, App
26	270	24.5	417	10	US-09-476-300-475	Sequence 475, App
27	270	24.5	417	14	US-10-017-754-475	Sequence 475, App
28	270	24.5	417	16	US-10-113-872-475	Sequence 475, App
29	270	24.5	417	17	US-10-283-017-475	Sequence 475, App
30	262	23.5	453	14	US-10-060-036-2622	Sequence 2622, App
31	259.4	23.5	451	14	US-10-066-543-1967	Sequence 1967, App
32	197.8	17.9	549	17	US-10-152-319A-124	Sequence 124, App
33	165	14.9	336	9	US-09-738-973-560	Sequence 560, App
34	165	14.9	336	9	US-09-854-133-560	Sequence 560, App
35	165	14.9	336	15	US-10-144-649A-560	Sequence 560, App
36	162.4	14.7	869	18	US-10-363-345A-29799	Sequence 29799, A
37	162.4	14.7	869	18	US-10-363-345A-29800	Sequence 29800, A
38	160.6	14.5	356	9	US-09-783-590-147	Sequence 147, App
39	149.8	13.6	869	18	US-10-363-345A-29797	Sequence 29797, A
40	149.8	13.6	869	18	US-10-363-345A-29798	Sequence 29798, A
41	141.4	12.8	474	10	US-09-918-995-26636	Sequence 26636, A
42	123.2	11.2	412	14	US-10-198-846-2573	Sequence 2573, App
43	118.8	10.8	550	17	US-10-362-674-1	Sequence 1, Appl1
44	118.8	10.8	1222	18	US-10-335-053-193	Sequence 193, App
45	115.8	10.5	433	9	US-09-983-965-2915	Sequence 2915, App

ALIGNMENTS

RESULT 1
US-10-071-645-1
; Sequence 1, Application US/10071645
; Publication No. US20030148389A1
; GENERAL INFORMATION:
; APPLICANT: Bihain, Bernard
; APPLICANT: Bour, Barbara
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Schizophrenia Related Gene and Protein
; FILE REFERENCE: 92.US2.CIP
; CURRENT APPLICATION NUMBER: US/10/071,645
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/223,482
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: PCR/IS01/01891
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 87..346
US-10-071-645-1

Query Match 100.0%; Score 1104; DB 15; Length 1104;
Best Local Similarity 100.0%; Pred. No. 1.6e-251;
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGAGCGCGCGCTGACCTGTCCGCGCGCGGGGACGCGCGCGGAGGCGG 60
DB 1 GGCACGAGCGCGCGCTGACCTGTCCGCGCGCGGGGACGCGCGCGGAGGCGG 60

```
QY 61 CCGCGCGGAGCCCGGACGCGACCATGTGGAGGTGCTGCCCTACGGCGACGAGC 120
Db 61 CCGCGCGGAGCCCGGACGCGACCATGTGGAGGTGCTGCCCTACGGCGACGAGC 120
QY 121 TGAGCCCTACGCGAGCGCGGCGAGCGTGGCCGAGATCTTCTCGCCGCTCGACGACA 180
Db 121 TGAGCCCTACGCGAGCGCGGCGAGCGTGGCCGAGATCTTCTCGCCGCTCGACGACA 180
QY 181 CCAACAACTTCTTTCGGCGCGCGGAGAAACGAGCGCGCGCCCAAGCTGGCGAGATCGGCC 240
Db 181 CCAACAACTTCTTTCGGCGCGCGGAGAAACGAGCGCGCGCCCAAGCTGGCGAGATCGGCC 240
QY 241 GAGCAAGCGGGTGTATTGAAGATCATAGGATTCATGACGCTGCTGAAATAATACACCG 300
Db 241 GAGCAAGCGGGTGTATTGAAGATCATAGGATTCATGACGCTGCTGAAATAATACACCG 300
QY 301 ACAAGGACCTCTGGTGTCTAACTCCCAAGACAATGAGTTAAGGAGAGATAGGAAC 360
Db 301 ACAAGGACCTCTGGTGTCTAACTCCCAAGACAATGAGTTAAGGAGAGATAGGAAC 360
QY 361 GCGGTAAAGTATTGGCAAAAGCATGAAAGAAAGCACTTTGAAATTTTACTA 420
Db 361 GCGGTAAAGTATTGGCAAAAGCATGAAAGAAAGCACTTTGAAATTTTACTA 420
QY 421 GCTTGTAACCCAGATGAATCAACACCTGTATCTGGTATATGCCCGGAGACAGATTAGG 480
Db 421 GCTTGTAACCCAGATGAATCAACACCTGTATCTGGTATATGCCCGGAGACAGATTAGG 480
QY 481 CGAAGGAGAGAGAGAGAGAGAAAGGCTTGGCCCTCTCAATAATAAATAAAAAA 540
Db 481 CGAAGGAGAGAGAGAGAGAGAAAGGCTTGGCCCTCTCAATAATAAATAAAAAA 540
QY 541 AATTTAAATAATAAATCCCTATATCCATATAGAAATGAAAGTCTCAGTCAGTA 600
Db 541 AATTTAAATAATAAATCCCTATATCCATATAGAAATGAAAGTCTCAGTCAGTA 600
QY 601 TTGGCAAAATTAATCCATTTCTTTTAAATACGGGAATATTGGCAATTATAGATCTGGATT 660
Db 601 TTGGCAAAATTAATCCATTTCTTTTAAATACGGGAATATTGGCAATTATAGATCTGGATT 660
QY 661 TTGACCACTTAATGAGCGGCGACCCAGGTGTTGAGGTGTTGGATCTTCGCTGATT 720
Db 661 TTGACCACTTAATGAGCGGCGACCCAGGTGTTGAGGTGTTGGATCTTCGCTGATT 720
QY 721 TGCGCTGTTCCCAATGTTTACATTTAATCTTGCAAAATGCTTCTGTGCACTTGGATG 780
Db 721 TGCGCTGTTCCCAATGTTTACATTTAATCTTGCAAAATGCTTCTGTGCACTTGGATG 780
QY 781 TGAATGCTGTCAGTTTATTTTATTTTATGTTGTTATCCTTGGATGTACAAAAATTCA 840
Db 781 TGAATGCTGTCAGTTTATTTTATTTTATGTTGTTATCCTTGGATGTACAAAAATTCA 840
QY 841 GAAATGATCTGTAGATATTCTGTTTATTTTATTTGTCATCTTTAGAGTTATCAGGAT 900
Db 841 GAAATGATCTGTAGATATTCTGTTTATTTTATTTGTCATCTTTAGAGTTATCAGGAT 900
QY 901 GTGTTTAAACAAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTATTTA 960
Db 901 GTGTTTAAACAAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTATTTA 960
QY 961 AATGAGTTGTAAGCTTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCA 1020
Db 961 AATGAGTTGTAAGCTTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCA 1020
QY 1021 AATGACATTTTTTATGTCAGAAAAACACACACACACACACACACACACACACAC 1080
Db 1021 AATGACATTTTTTATGTCAGAAAAACACACACACACACACACACACACACACAC 1080
QY 1081 CACGAAAAAATAAAAAA 1104
Db 1081 CACGAAAAAATAAAAAA 1104
```

```
RESULT 2
US-09-820-649-25
; Sequence 25, Application US/09820649
; Publication No. US20030199683A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 83 Human Secreted Proteins
; FILE REFERENCE: P2012PI
; CURRENT APPLICATION NUMBER: US/09/820,649
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US/09/236,557
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: PCT/US98/15949
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,212
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,209
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,234
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,218
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,214
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,236
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,215
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,211
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (39)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-820-649-25
```

```
Query Match 63.4%; Score 699.8; DB 10; Length 1510;
Best Local Similarity 89.7%; Pred. No. 1.7e-155;
Matches 893; Conservative 1; Mismatches 69; Indels 33; Gaps 12;

QY 141 GGGCAGCTGGGCCAGATCTTCTCTCGCGCTCGAGACACCAACACTTCTTCGGCGCC 200
Db 2 GGGCAGCTGGGCCAGATCTTCTCTCGCGCTCGAGACACCAACACTTCTTCGGCGCC 61
QY 201 GGGCAGAACAGCGCGCCCAAGCTGGCCAGATCGCCGGAGCAAGCGGTTGTTATT 260
Db 62 GGGCAGAACAGCGCGCCG-CCAAGCTGGCCAGATCGCCGGAGCAAGCGGTTGTTATT 120
QY 261 GAAGATGATAGATTGATGACGTCTGAAAAATATGACCGACAGGACCT-CTGGTGTG 319
Db 121 GAAGATGATAGATTGATGACGTCTGAAAAATATGACCGACAGGACCTCTTCGGTGTG 180
QY 320 TAACT-CCCCAAAGACAATGAGTTAAGGAGAGAAATAGGAAACGGCGTAAACAGTTATTGG 378
Db 181 TAACTCCCCAAAGACAATGAGTTAAGGAGAGAAATAGGAAACGGCGTAAACAGTTATTGG 240
QY 379 CAAAAGCATGAAAGAGAAAGCACTTTGAAATTTTATTACTAGCTTG-TACCCACGATGA 437
Db 241 CAAAAGCATGAAAGAGAAAGCACTTTGAAATTTTATTACTAGCTTGCTACCCACGATGA 300
QY 438 AATCAACAACCTGTATCTGTAT-ATGCCGGAGACAGATTAGGCGA-AGGAGGAAGAGA 495
Db 301 AATCAACAACCTGTATCTGTATCAGGCCGGGAGACAGATGAGCGAGAGAGGAGG 360
QY 496 GAGAGAAAGAGGCTTGGGC-CCTCTACAAATAAATAAAAAAATAAATAAATAAATAA 554
```



```
QY 1009 CAAAGTAATGCAATGGACACATTTTATGTCAGAAAAACACACACACACACACACA 1068
Db 901 CAAAGTAATGCAATGGACACATTTTATGTCAGAAAAACACACACACACACACACA 960
QY 1069 CACACACACACACACACACACACACACACACACACACACACACACACACACAC 1104
Db 961 CACACACACACACACACACACACACACACACACACACACACACACACACACAC 996

RESULT 5
US-10-723-860-7058/c
; Sequence 7058, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NFUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7058
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1952)..(1952)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2033)..(2033)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7058

Query Match 58.2%; Score 642; DB 18; Length 2058;
Best Local Similarity 87.3%; Pred. No. 1.le-141;
Matches 838; Conservative 0; Mismatches 90; Indels 32; Gaps 11;

QY 177 GACACCAACAACCTCTCGCGCGCGGCGGACACACACACACACACACACACACAC 236
Db 1873 GAGACAAACCTCGATCGGCTGCGCGCGGCGGACACACACACACACACACACAC 1814

QY 237 GCGCGGAGCAAGCGGCTGTTATTGAAGATGATGATGATGATGATGATGATGATG 296
Db 1813 GCGCGGAGCAAGCGGCTGTTATTGAAGATGATGATGATGATGATGATGATGATG 1754

QY 297 ACCGACAGGCAACCT-CTGGTGTCTAACT-CCCCAAGACATGATGATGATGATGAT 354
Db 1753 ACCGACAGGCAACCTCTGGTGTCTAACTCCCCAAGACATGATGATGATGATGATG 1694

QY 355 AGGACAGCGGTAACAGTATTGGCAAAAGATGATGATGATGATGATGATGATGATG 414
Db 1693 AGAAGCGGCGTAACAGTATTGGCAAAAGATGATGATGATGATGATGATGATGATG 1634

QY 415 TTACTAGCTTG-TACCCAGATGAATCAACACCTGTATCTGGTAT-ATGCCCGGAGAC 472
Db 1633 TTACTAGCTTG-TACCCAGATGAATCAACACCTGTATCTGGTATCTGGTATCTGG 1574

QY 473 AGATTAGGCGA-AGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 530
Db 1573 AGATTAGGCGA-AGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1514

QY 531 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 574
Db 1513 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1454

QY 575 AAGAATAAAGAGTCTCACT- - - -GCAGTATTGGCAAAATTAATCAATTCATTTT 630
```

```
Db 1453 AAATAAAGAAAGTCTCAGTATTGTCAGAAATTAATATCCATTTCTTTTATA 1394
QY 631 ACGGGAATATT- - - -GGCATTATACATCTGGATTTTG-ACCACTTAATGAAGCGG-ACC 684
Db 1393 TACGGTGAATATTGGCAATATATAGATCTGGATTTTGAAACCACTTANTGNAGCGGCAACA 1334
QY 685 CCAGGTGTTTTGAGGTGTGGCATTCTTTGCGTGATTTGGCTTTCCCAATGTTTACATTA 744
Db 1333 CCAGGTGTTTTGAGGTGTGGCATTCTTTGCGTGATTTGGCTTTCCCAATGTTTACATTA 1274
QY 745 TTTAATCTTGCAAAATGTTCTGTCACCTTGGATGTGAAATGCTGTCAGTTTATTTT 804
Db 1273 TTTAATCTTGCAAAATGTTCTGTCACCTTGGATGTGAAATGCTGTCAGTTTATTTT 1214
QY 805 TTTTATGTTGTTATCTCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATTTCT 864
Db 1213 TTTTATGTTGTTATCTCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATTTCT 1154
QY 865 GTTTTATTTTGGTCACTCTTTAGAGTTATCAGGAATGTGTTTAAACCAAGAGAGAACTT 924
Db 1153 GTTTTATTTTGGTCACTCTTTAGAGTTATCAGGAATGTGTTTAAACCAAGAGAGAACTT 1094
QY 925 TTCTAAGGAATGATACATAGAAAGATTTTATTTTAAATGATGTTGTTAAAGCTTGTGTTT 984
Db 1093 TTCTAAGGAATGATACATAGAAAGATTTTATTTTAAATGATGTTGTTAAAGCTTGTGTTT 1034
QY 985 CTTTGTGCTGCAAGCTATCTGCCCAAGTTAATGCAATGGAACACATTTTATGTCACA 1044
Db 1033 CTTTGTGCTGCAAGCTATCTGCCCAAGTTAATGCAATGGAACACATTTTATGTCACA 974
QY 1045 AAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 1104
Db 973 AAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 914

RESULT 6
US-10-198-846-11350
; Sequence 11350, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11350
; LENGTH: 1850
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 26, 107, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846,
; LOCATION: 1847, 1848, 1849, 1850
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11350

Query Match 57.5%; Score 634.6; DB 14; Length 1850;
Best Local Similarity 86.9%; Pred. No. 5.9e-140;
Matches 833; Conservative 0; Mismatches 94; Indels 32; Gaps 11;

QY 177 GACACCAACAACCTTTCTGGCGCGCGGCGGACACACACACACACACACACACAC 236
Db 192 GAAACCTCGATCGGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 251
```

```
QY 237 GCGCGGAGCAAGCGGTTCTTATGAAGATAGATAGATTGATGCTGCTGAAATATG 296
Db 252 GCGCGGAGCAAGCGGTTCTTATGAAGATAGATAGATTGATGCTGCTGAAATATG 311
QY 297 ACCGACAAGGCACTTCTGCTCTAACT-CCCCAAGACAATGATGTTAAGGAGAGAAAT 354
Db 312 ACCGACAAGGCACTTCTGCTCTAACTCCCCAAGACAATGATGTTAAGGAGAGAAAT 371
QY 355 AGGAACGGCGGTAAACAGTTATTTGGCAAAAGATGAAAGAGAAAGCACTTTGAAATTTA 414
Db 372 AAGAACGGCGGTAAACAGTTATTTGGCAAAAGATGAAAGAGAAAGCACTTTGAAATTTA 431
QY 415 TTACTAGCTTG-TACCCAGATGAATCAACACCTGTATCTGCTAT-ATGCCGCGAGAC 472
Db 432 TTACTAGCTTGTTACCCAGATGAATCAACACCTGTATCTGCTATCTGCTGCGCGGAGAC 491
QY 473 AGATTAGGCGA-AGGAGGAGAGAGAGAGAGAAAGGCTTGGGC-CCTCTACAAATAAAA 530
Db 492 AGATGAGCGGAGAGGAGGAGGAGGAGAGAGAGGCTCTGGGCTCTCTGCAAAATAA 551
QY 531 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 574
Db 552 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 611
QY 575 AAGAATAAAGAGTCTCAGT---GCAGTATTTGGCAAAATAAATAAATAAATAAATAA 630
Db 612 AATAAATAAAGAGTCTCAGTTCGAGCTATTTGTCAAAATAAATAAATAAATAAATAA 671
QY 631 ACGGGAATATT---GGCATTATAGATCTGGAATTTTG-ACCATTAAATGAAGCGC-ACC 684
Db 672 TACGGTGAATATTGGCAATTTATAGATCTGGATTTTGAACCACTTAATGAAGCGCAACA 731
QY 685 CAGGTGTTTGGAGTGTGGCAATCTTCGCTGATTTGCTGCTGTTGCCAATGTTTACATTA 744
Db 732 CCAGGTGTTTGGAGTGTGGCAATCTTCGCTGATTTGCTGCTGTTGCCAATGTTTACATTA 791
QY 745 TTTAATCTTGAATAAGTCTCTGCTGATGTAAGTGTGAAATGCTGCTGCAAGTTTATTTT 804
Db 792 TTTAATCTTGAATAAGTCTCTGCTGATGTAAGTGTGAAATGCTGCTGCAAGTTTATTTT 851
QY 805 TTTTATGTTGTTATCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATCT 864
Db 852 TTTTATGTTGTTATCCTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATCT 911
QY 865 GTTTATTTTGGTCATCTTTAGAGTTATCAGGAATGTTTAAACAGAGAGACTT 924
Db 912 GTTTATTTTGGTCATCTTTAGAGTTATCAGGAATGTTTAAACAGAGAGACTT 971
QY 925 TTCTAAGGAATGATACATAGAAAAAGTTTATTTTAAATGAGTTGTAAGCTTGTGTTT 984
Db 972 TTCTAAGGAATGATACATAGAAAAAGTTTATTTTAAATGAGTTGTAAGCTTGTGTTT 1031
QY 985 CTTTGTGTCGAAGCTATCTGCCAAGTTAATGCAAAATGGAACACATTTTTTATGTCAGA 1044
Db 1032 CTTTGTGTCGAAGCTATCTGCCAAGTTAATGCAAAATGGAACACATTTTTTATGTCAGA 1091
QY 1045 AAAAAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 1103
Db 1092 AAAAAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 1150
```

RESULT 7
US-10-172-118-1796
; Sequence 1796, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura

; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1796
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 018584
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1796

Query Match 57.3%; Score 632.6; DB 17; Length 1321;
Best Local Similarity 89.0%; Pred. No. 1.5e-139;
Matches 816; Conservative 0; Mismatches 69; Indels 32; Gaps 11;
QY 220 CCAAGCTGGCGCAGATCGGCGGAGCAAGCGGTTGTTATTGAAGATAGTAGATTGATG 279
Db 1 CCAAGCTGGCGCAGATCGGCGGAGCAAGCGGTTGTTATTGAAGATAGTAGATTGATG 60
QY 280 ACGTGCTGAAAAATATGACCGCAAGGCACT-CTGGTGTCTAACT-CCCCAAGACAAT 337
Db 61 ACGTCTGAAAAATATGACCGCAAGGCACTCTCTGGTGTCTAACTCCCCAAGACAAT 120
QY 338 GAGTTAAGGAGAGAAATAGGAACCGCGGTAACTATTGGCAAAAGCATGAAAGAGA 397
Db 121 GAGTTAAGGAGAGAAATAGGAACCGCGGTAACTATTGGCAAAAGCATGAAAGAGA 180
QY 398 AAGCACTTTGAAATTTATTACTAGCTTG-TACCCAGATGAATCAACAACTGTATCTG 456
Db 181 AAGCACTTTGAAATTTATTACTAGCTTGCTACCCAGATGAATCAACAACTGTATCTG 240
QY 457 GTAT-ATGCCCGGAGACAGATTAGCGA-AGGAGAGAGAGAGAGAGAGAGGCTTGGG 514
Db 241 GTATCAGCGCGGAGACAGATGAGCGAGAGAGAGAGAGAGAGAGGAGGAGGCTCTGGG 300
QY 515 C-CTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 566
Db 301 CTCTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 360
QY 567 -----TCCCATATAAAGATTAAGACGCTCAGT-----GCAGTATGGCAAAATTA 613
Db 361 TATATACATATAAAGAAATAAAGAAAGTCTCAGTTGCGAGCTATTTGTCAAAATTAAT 420
QY 614 ATCCATTTCTTTTAAATACGGG---AATATTGGCATTATAGATCTGGAATTTG-ACCAC 668
Db 421 ATCCATTTCTTTTATATACGGTGAATATGCGCAATTAAGATCTGATTTTGAACCA 480
QY 669 TTAATGAAGCGGC-ACCCAGGTGTTTGGGTGTGGCAATCTTTCGCTGATTTGCGTGT 727
Db 481 TTAATGAAGCGGCACACACAGGTGTTTGGGTGTGGCAATCTTTCGCTGATTTGCGTGT 540
QY 728 TCCCAATGTTTACATTTTAACTCTGCAAAATGGTCTGTGCACTTGGATGTGAATG 787
Db 541 TCCCAATGTTTACATTTTAACTCTGCAAAATGGTCTGTGCACTTGGATGTGAATG 600
QY 788 CTGTCAGTTTATTTTATTTTATGTTATCTCTGGATGTACAAAAATTCAGAAAAATG 847
Db 601 CTGTCAGTTTATTTTATTTTATGTTATCTCTGGATGTACAAAAATTCAGAAAAATG 660
QY 848 ATCTCTGTAGATATCTGTTTATTTTCTGTCATCTTTAGAGTTATCAGGAATGTGTTTA 907
Db 661 ATCTCTGTAGATATCTGTTTATTTTCTGTCATCTTTAGAGTTATCAGGAATGTGTTTA 720
QY 908 AACAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAGATTTTATTTTAAATGAG 967
Db 721 AACAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAGATTTTATTTTAAATGAG 780

Qy	968	TTGTAAAGCTTGTGTTCTTTGTTGTTGCTGCAAGCTATCTGCCCAGTAATGCAAAATGGAC	1027
Db	781	TTGTAAAGCGTGTGTTCTTTGTTGTTGCTGCAAGCTATCTGCCCAGTAATGCAAAATGGAC	840
Qy	1028	ACATTTTTTTATGTGCAGAAAAACACACACACACACACACACACACACACACACACACAC	1087
Db	841	ACATTTTTTTATGTGCAGAAAAACACACACACACACACACACACACACACACACACACAC	900
Qy	1088	AAAAAAAAAAAAAAAAAAAA 1104	
Db	901	ACGAAAAACCAAGAAAA 917	

RESULT 8

US-10-342-887-1796

; Sequence 1796, Application US/10342887

; Publication No. US20040058340A1

; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue

; APPLICANT: He, Yudong

; APPLICANT: Linsley, Peter S.

; APPLICANT: Mao, Mao

; APPLICANT: Roberts, Christopher J.

; APPLICANT: Van 't Veer, Laura Johanna

; APPLICANT: Van de Vijver, Marc J.

; APPLICANT: Bernards, Rene

; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-188-999

; CURRENT APPLICATION NUMBER: US/10/342,887

; CURRENT FILING DATE: 2003-01-15

; PRIOR APPLICATION NUMBER: 60/298,918

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: 60/380,710

; PRIOR FILING DATE: 2002-05-14

; PRIOR APPLICATION NUMBER: 10/172,118

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 2699

; SEQ ID NO 1796

; LENGTH: 1321

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-342-887-1796

```

361 Db TATATACACATATAAGAAATAAAAGAAAGTCTCAGTTGCAGCTATTTTGTCAAAATTAAT 420
Qy 614 ATCCATTTCTTTTAAATACGGG----AATATTGGCATTATAGATCTGGATTTTG-ACCAC 668
Db 421 ATCCATTTCTTTTATATACGGTGAATATTGGCAATTATAGATCTGGATTTTGAACCAC 480
Qy 669 TTAATGAAGCGGC-ACCCACAGGTGTTTGAAGGTGTGGCAATCTTCGCTGATTTGGCTGT 727
Db 481 TTAATGAAGCGGCAACACACAGGTGTTTGAAGGTGTGGCAATCTTCGCTGATTTGGCTGT 540
Qy 728 TCCCAATGTTTACATTAATTTAATCTTGCAGAAATAGTTCTGTGCACCTTGGATGTGAATG 787
Db 541 TCCCAATGTTTACATTAATTTAATCTTGCAGAAATAGTTCTGTGCACCTTGGATGTGAATG 600
Qy 788 CTGTCCAGTTTTATTTTTTTTATGTTGTTATCTGTGATGTACAAAAAATTCAGAAAAATG 847
Db 601 CTGTCCAGTTTTATTTTTTTTATGTTGTTATCTTGTGTTATCTTTAGAGTTATCAGGAATG 660
Qy 848 ATCTCTGTAGATATCTTGTTTTATTTTTTGGTTCATCTTTAGAAAGTTATCAGGAATGTGTTTA 907
Db 661 ATCTCTGTAGATATCTTGTTTTATTTTTTGGTTCATCTTTAGAGTTATCAGGAATGTGTTTA 720
Qy 908 AAACAAGAGAGAACTTTTCTTAAGGAATGATACATAGAAAAAGATTTTATTTAAAAATGAG 967
Db 721 AAACAAGAGAGAACTTTTCTTAAGGAATGATACATAGAAAAAGATTTTATTTAAAAATGAG 780
Qy 968 TTGTAAGCTTGTGTTCTTTGTTGTTGTCGCAAGCTATCTGCCCAAGTTAATGCAANTGGAC 1027
Db 781 TTGTAAGCGTGTGTTCTTTGTTGTCGCAAGCTATCTGCCCAAGTTAATGCAANTGGAC 840
Qy 1028 ACATTTTTTATGTCAGAAAAACACACACACACACACACACACACACACACACACACACAC 1087
Db 841 ACATTTTTTATGTCAGAAAAACACACACACACACACACACACACACACACACACACACAC 900
Qy 1088 AAAAAAAAAAAAAAAAAA 1104
Db 901 ACGAAAAACAAGAAAA 917

RESULT 9
US-10-071-645-3
; Sequence 3, Application US/10071645
; Publication No. US20030148389A1
; GENERAL INFORMATION:
; APPLICANT: Bihaun, Bernard
; APPLICANT: Bour, Barbara
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: Schizophrenia Related Gene and Protein
; FILE REFERENCE: 92.US2.CIP
; CURRENT APPLICATION NUMBER: US/10/071,645
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/223,482
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01891
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent.pm
; SEQ ID NO 3
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-071-645-3

```

Query Match	55.0%	Score	607.2;	DB	15;	Length	3189;
Best Local Similarity	89.2%;	Pred. No.	2.5e-133;				
Matches	787;	Conservative	0;	Mismatches	63;	Indels	32;
Gaps	11;						

Qy	252	GTGTGTTATTGAAGATGATAGGATTGATGACGTC	CTGAAAAATATGACCGCAAGGCACCT	311
Db	2158	GTGTGTTATTGAAGATGATAGGATTGATGACGTC	CTGAAAAATATGACCGCAAGGCACCT	2217
Qy	312	-CTGGTGTCCTAACT-CCCCAAGACAAATGAGTTAAGGGAGAGAAATAGGAAACGCGCGTAAAC	369	

Db 361 ATTGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTATTACTAGCTTGTCTACCCAC 420
QY 433 GATGAATCAACAACCTGTATCTGGTAT-ATGCCGGGACAGATTTAGCGGA-AGGAGGA 490
Db 421 GATGAATCAACAACCTGTATCTGGTATCAGCGCGGACAGATCAGCGGAGAGGA 480
QY 491 AGAGAGAGAGAAAGGCTTGGGCCCTCTACAATATAAATAAAAAAAAAATTTAAA 550
Db 481 CGAGGAGGAGGAGAGGCTCTGGGGCTCTCTGCAAAAAAAAAAAAAAAAAAAAAA 540
QY 551 TAATAAAA 558
Db 541 AAAAAAAA 548

RESULT 11
US-09-978-360A-155
; Sequence 155, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 155
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 327..416
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 404..417
US-09-978-360A-155

Query Match 36.7%; Score 405.4; DB 11; Length 417;
Best Local Similarity 98.6%; Pred. No. 6.6e-86;
Matches 409; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 690 TGTGTTGAGGTGTCGATTTCTGCTGATTTGGCTGTTCCCAATGTTTACATTTTAA 749
Db 1 TGTGTTGAGGTGTCGATTTCTGCTGATTTGGCTGTTCCCAATGTTTACATTTTAA 60
QY 750 TCTTGCAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGCCAGTTTATTTTTTTA 809

Db 61 TCTTGCAAAATGGTTCTGTGCACTTGGATGTGAAATGCTGCCAGTTTATTTTTTA 120
QY 810 TGTGTTATCTTGTGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTTCTGTTT 869
Db 131 TGTGTTATCTTGTGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTTCTGTTT 180
QY 870 ATTTTGGTCATCTTTTGAAGTTTATCAGGAATGTGTTTAAAAACAAGAAGAACTTTTCTA 929
Db 181 ATTTTGGTCATCTTTTGAAGTTTATCAGGAATGTGTTTAAAAACAAGAAGAACTTTTCTA 240
QY 930 AGGAATGATACATAGAAAAAGATTTTATTTTAAAAATGAGTTGTAAGCTTGTCTTCTT 989
Db 241 AGGAATGATACATAGAAAAAGATTTTATTTTAAAAATGAGTTGTAAGCTTGTCTTCTT 300
QY 990 TTGCTGCAAGCTATCTGCCCAAGTTTAAATGCAAAATGGACACACATTTTATGTGCAAAAAAC 1049
Db 301 TTGCTGCAAGCTATCTGCCCAAGTTTAAATGCAAAATGGACACACATTTTATGTGCAAAAAAC 360
QY 1050 AC 1104
Db 361 ACACACACACACACACACACATATACACACACACACACACACACACACACACACACACAC 415

RESULT 12
US-10-027-632-231746
; Sequence 231746, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 231746
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-231746

Query Match 34.5%; Score 380.4; DB 13; Length 584;
Best Local Similarity 98.0%; Pred. No. 6.8e-80;
Matches 384; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 713 CGCTGATTTGGCTGTCTCCCAATGTTTACATTTTAAATCTGCAAAAAATGGTTCTGTGCA 772
Db 1 CGCTGATTTGGCTGTCTCCCAATGTTTACATTTTAAATCTGCAAAAAATGGTTCTGTGCA 60
QY 773 CTTGATGTAAGATGCTGCTCCAGTTTATTTTTTATGTTGTTATCTTGGATGTACAA 832
Db 61 CTTGATGTAAGATGCTGCTCCAGTTTATTTTTTATGTTGTTATCTTGGATGTACAA 120

QY 1091 AAAAAA 1097
|||||
Db 360 AAAACA 366

RESULT 15

US-10-085-783A-49104
; Sequence 49104, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49104
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-49104

Query Match 31.1%; Score 343.8; DB 17; Length 366;
Best Local Similarity 97.8%; Pred. No. 2.6e-71;
Matches 359; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 731 CAATCTTACATATTAATCTTGCAAAATGGTCTGCGACTTGGATGCGAATGCTG 790
Db 1 CAATGTTTACATATTAATCTTGCAAAATGGTCTGCGACTTGGATGCGAATGCTG 60

QY 791 TCCAGTTTATTTTATTTTATGTTTATCTCTTGGATGTACAAAAATTCAGAAAATGATC 850
Db 61 TCCAGTTTATTTTATTTTATGTTTATCTCTTGGATGTACAAAAATTCAGAAAATGATC 120

QY 851 TCTGTAGATATTCGTGTTTATTTTGGTCACTTTTGAAGTTATCAGGAATGTGTTAAAA 910
Db 121 TCTGTAGATATTCGTGTTTATTTTGG-CATCTTTAGAAGTTATCAGGAATGTGTTAAAA 179

QY 911 CAAGAAGAGAACTTTTCTAAGGAATCATACATAGAAAAGATTTTATTTAAATGAGTTG 970
Db 180 CAAGAAGAGAACTTTTCTAAGGAATCATACATAGAAAAGATTTTATTTAAATGAGTTG 239

QY 971 TAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAAATGGACACA 1030
Db 240 TAAAGCTTGTGTTTCTTTGTTGCTGCAAGCTATCTGCCCAAGTTAATGCAAAATGGACACA 299

QY 1031 TTTTATGTCAGAAAAAC 1090
Db 300 TTTTATGTCAGAAAAAC 359

QY 1091 AAAAAA 1097
|||||
Db 360 AAAACA 366

Search completed: March 30, 2005, 05:50:57
Job time : 670 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2005, 18:29:28 ; Search time 3513 Seconds
(without alignments)
11962.127 Million cell updates/sec

Title: US-10-071-645-1
Perfect score: 1104
Sequence: 1 ggcacgagcgagcgccgctg.....aaaaaaaaaaaaaaaaaaaa 1104

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	632.6	57.3	1321	3 AF116637	AF116637 Homo sapi
2	623	56.4	954	7 CN643673	CN643673 ILLUMIGEN
3	615.6	55.8	945	7 CN803102	CN803102 ILLUMIGEN
4	614.4	55.7	976	7 CN802133	CN802133 ILLUMIGEN
5	611.8	55.4	953	7 CN803224	CN803224 ILLUMIGEN
6	611.2	55.4	950	7 CN646790	CN646790 ILLUMIGEN
7	610.8	55.3	954	7 CN646555	CN646555 ILLUMIGEN
8	610.2	55.3	954	7 CN646445	CN646445 ILLUMIGEN
9	609.6	55.2	952	7 CN646545	CN646545 ILLUMIGEN
10	608.2	55.1	951	7 CN647038	CN647038 ILLUMIGEN
11	607.2	55.0	1566	3 CR604926	CR604926 full-leng
12	606.4	54.9	955	7 CN644007	CN644007 ILLUMIGEN
13	605.4	54.8	954	7 CN803392	CN803392 ILLUMIGEN
14	599.6	54.3	953	7 CN646019	CN646019 ILLUMIGEN
15	589.8	53.4	964	7 CN646785	CN646785 ILLUMIGEN
16	586.6	53.1	955	7 CN805169	CN805169 ILLUMIGEN
17	581.6	52.7	955	7 CN802272	CN802272 ILLUMIGEN
18	580.2	52.6	951	7 CN802081	CN802081 ILLUMIGEN
19	579.8	52.5	1020	7 CN803473	CN803473 ILLUMIGEN
20	574.8	52.1	944	7 CN803534	CN803534 ILLUMIGEN
21	574.4	52.0	955	7 CN801954	CN801954 ILLUMIGEN
22	557.2	50.5	1026	7 CN643857	CN643857 ILLUMIGEN
23	550.6	49.9	993	7 CN802904	CN802904 ILLUMIGEN
24	548.8	49.7	956	7 CN805265	CN805265 ILLUMIGEN

25	542.8	49.2	911	7	CN802110	CN802110 ILLUMIGEN
26	532.6	48.2	967	7	CN644185	CN644185 ILLUMIGEN
27	520.8	47.2	967	7	CN641685	CN641685 ILLUMIGEN
28	515.2	46.7	968	7	CN805606	CN805606 ILLUMIGEN
29	511.2	46.3	806	7	CN642220	CN642220 ILLUMIGEN
30	503.4	45.6	1001	7	CN802311	CN802311 ILLUMIGEN
31	500.4	45.3	1004	7	CN801895	CN801895 ILLUMIGEN
32	499.6	45.3	983	7	CN803331	CN803331 ILLUMIGEN
33	496	44.9	890	7	CO645235	CO645235 ILLUMIGEN
C 34	489.6	44.3	747	5	BQ018543	BQ018543 UI-H-DH1-
35	483	43.8	981	7	CN801823	CN801823 ILLUMIGEN
36	471.2	42.7	914	7	CO645211	CO645211 ILLUMIGEN
37	463.2	42.0	819	4	BI553632	BI553632 603190553
38	447.8	40.6	953	7	CO645289	CO645289 ILLUMIGEN
C 39	446	40.4	597	7	W17400	W17400 zbl5b10.r1
C 40	445.2	40.3	661	5	BQ000750	BQ000750 UI-H-DH1-
41	444.2	40.2	1083	3	CR606015	CR606015 full-leng
C 42	441	39.9	676	5	BQ771726	BQ771726 UI-H-E21-
43	440.6	39.9	931	5	EX419572	EX419572 BX419572
44	437.4	39.6	695	7	CN428017	CN428017 170006000
45	435.8	39.5	808	5	BU570962	BU570962 AGENCOURT

ALIGNMENTS

RESULT 1
AF116637/c 1321 bp mRNA linear HTC 08-MAY-2001

LOCUS AF116637 Homo sapiens PRO1489 mRNA, complete cds.
DEFINITION AF116637
ACCESSION AF116637
VERSION AF116637.1 GI:7959775
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1321)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhou, G., Ouyang, S., Luo, L., Bi, J., Liu, M. and He, F.
TITLE Functional prediction of the coding sequences of 121 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1321)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhou, G., Ouyang, S., Luo, L., Bi, J., Liu, M. and He, F.
TITLE Direct Submission
JOURNAL Submitted (24-DEC-1998) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing 100850, P. R. China
FEATURES
source Location/Qualifiers
1..1321
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FLB5706"
/tissue_type="liver"
/dev_stage="fetal"
complement(991..1152)
note="predicted protein of HQ1489"
/codon_start=1
/product="PRO1489"
/protein_id="AAF71060.1"
/db_xref="GI:7959775"
/translation="MKRESTLKFTSLPTMKSTTCIWIYQAGRQRREBEKEKALGSSAKIKKK"

CDS
Query Match 57.3%; Score 632.6; DB 3; Length 1321;
Best Local Similarity 89.0%; Pred. No. 5.5e-124;
Matches 816; Conservative 0; Mismatches 69; Indels 32; Gaps 11;

ORIGIN
220 CCAAGCTGGCCAGATCGCCGAGCAAGCGGTTGTTATTGAAGATGATAGGATTGATG 279
SSAKIKKK

1321	Db	CCAGCTGGGCCAGATCGGCCGAGCAAGCGGCTGTATTGTAAGATGATAGGATTTGATG	1262
280	Qy	ACGTGCTGAAAAATATGACCGCAAGAAGCAGCT--CTGGTGTCTAACT--CCCAAGACAAT	337
1261	Db	ACGTGCTGAAAAATATGACCGCAAGAAGCAGCTCTCTGGTGTCTAACTCTCCCAAGACAAT	1202
338	Qy	GAGTTAAGGAGAGAATAGGAAACGGCGGTAAACAGTTATTGGCAAAAAGCATGAAAAGAGA	397
1201	Db	GAGTTAAGGAGAGAATAGGAAACGGCGGTAAACAGTTATTGGCAAAAAGCATGAAAAGAGA	1142
398	Qy	AAGCATTCTGAAATTTATTACTAGCTTG--TACCCACGATGAATCAACACCTGTATCTG	456
1141	Db	AAGCATTCTGAAATTTATTACTAGCTTGCTTACCCACGATGAATCAACACCTGTATCTG	1082
457	Qy	GTAT-ATGCCCGGAGACAGATTAGGCGA--AGGAGGAAGACAGAGAGAAGAAAGCCTTGGG	514
1081	Db	GTATCAGGCCGGGAGACAGATGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGCTCTGGG	1022
515	Qy	C-CCTCTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	566
1021	Db	CTCCTCTGCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	962
567	Qy	-----TCCCATATAAGAAATAAAGAGTCTCAGT---GCAGTATTGGCAAAATTA	613
961	Db	TATATACACATATAAGAAATAAATAAAGAGTCTCAGTTGCGAGCTATTGTCAAAATTAAT	902
614	Qy	ATCCATTCTCTTTTAATACGGG---AATATTGGCAATTATAGATCTGGAATTTTG-ACCAC	668
901	Db	ATCCATTCTCTTTTATATACGGGTGAATATTGGCAATTATAGATCTGGAATTTTGAACCCAC	842
669	Qy	TTAATCAAGCGGC-ACCCACAGGTGTTTTGAGGTGTGGCAATCTTCGCTGATTTGGCTGT	727
841	Db	TTAATCAAGCGGCACACACAGGTGTTTGAAGGTGTGGCAATCTTCGCTGATTTGGCTGT	782
728	Qy	TCCCAATGTTTACATTAATTAATCTTGCAAAAATGTTCTGTGCATCTGGATGTGAAATG	787
781	Db	TCCCAATGTTTACATTAATCTTGCAAAAATGTTCTGTGCATCTGGATGTGAAATG	722
788	Qy	CTGTCCAGTTTATTTTTTTATGTTGTTATCCTTGGATGTACAAAAATTCAGAAAAATG	847
721	Db	CTGTCCAGTTTATTTTTTTATGTTGTTATCCTTGGATGTACAAAAATTCAGAAAAATG	662
848	Qy	ATCTCTGTAGATATCTGTTTTTATTTTTTGGTCACTCTTTTAGAAGTTATCAGGAATGCTTTTA	907
661	Db	ATCTCTGTAGATATCTGTTTTTATTTTTTGGTCACTCTTTTAGAAGTTATCAGGAATGCTTTTA	602
908	Qy	AAACAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTTTAAAAATGAG	967
601	Db	AAACAAGAGAGAACTTTTCTAAGGAATGATACATAGAAAAAGATTTTTTAAAAATGAG	542
968	Qy	TTGTAAAGCTTGTTTTCTTTGTTGTGTCAGCTATCTGCCNAGTTTATGCAATGGAC	1027
541	Db	TTGTAAAGCGTGTGTTTCTTTGTTGTGTCAGCTATCTGCCNAGTTTATGCAATGGAC	482
1028	Qy	ACATTTTATGTCAGAAAAACACACACACACACACACACACACACACACACACACGAAA	1087
481	Db	ACATTTTATGTCAGAAAAACACACACACACACACACACACACACACACACACACAC	422
1088	Qy	AAAAAAAAAAAAAAAAA 1104	
421	Db	ACGAAAAACAAGAAAA 405	

RESULT 2
CN643673
LOCUS
DEFINITION
ILLUMIGEN MCQ_8340 Katze MMR Macaca mulatta cDNA clone IBIUW:3286
5', similar to Bases 1 to 360 highly similar to human CAMK1Aalpha
(Hs.197922), mRNA sequence.
CN643673
CN643673
ACCESSION
CN643673.1
VERSION
KEYWORDS
GI:47154683
EST.
linear
954 bp
EST 12-MAY-2004

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA

Tel: 2063780400

Fax: 2063780408

Email: cmagness@illumigen.com

Sequenced on 2004.04.02. 813 Q20 bases.

PCR Primers

FORWARD: CCTCACTAAAGGGAACAAA

BACKWARD: CACTATAGGCGCAATTGGGTA

Insert Length: 953 Std Error: 0.00

Plate: CL000288 row: A column: 11

Seq primer: CCTCACTAAAGGGAACAAA

POLYA=yes.

FEATURES

source

Location/Qualifiers

1..953

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

/clone="IBIUM:13763"

/sex="female"

/dev_stage="adult"

/lab_host="E. coli SOLR"

/clone_lib="Katze MMR"

/note="Organ: Brain; Vector: Uni-ZAP XR; Site 1: Ecor I; Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)"

ORIGIN

Query Match 55.4%; Score 611.8; DB 7; Length 953;

Best Local Similarity 86.5%; Pred. No. 1.4e-119;

Matches 824; Conservative 0; Mismatches 92; Indels 37; Gaps 12;

```

QY 185 CAACCTCTCGGCGCGGCGGACAGCGCGGCCAGCTGGGCCAGATCGGCGGAG 244
DB 1 CAACCTCTCTCGGCGCGGCGGACAGCGCGGCCAGCTGGGCCAGATCGGCGGAG 60

QY 245 CAAGCGGGTGTATTGAAGATGATAGGATTGATGACGTGCTGCAAAAATATGACCGACAA 304
DB 61 CAAGCGGGTGTATTGAAGATGATAGGATTGATGACGTGCTGCAAAAATATGACCGACAA 120

QY 305 GGCACCT-CTGTGTCCTAACT-CCCCAAAGACAATGAGTTAAGGGAGAGAAATAGAACGG 362
DB 121 GGCACCTCTCGTGTCTAACTCCCCAAAGACAATGAGTTAAGGGAGAGAAATAGAACGG 180

QY 363 CGGTACACGTTATTCGCAAAAGCATGAAAGGAAAGCACATTTGAAATTTTACTAGC 422
DB 181 CGATACACGTTATTCGCAAAAGCATGAAAGGAAAGCACATTTGAAATTTTACTAGC 240

QY 423 TTG-TACCCACGATGAAATCAACCACTGTATCTGGTAT-ATGCCCGGAGACAGATTAGG 480
DB 241 TTGCTACCCACGATGAAATCAACCACTGTATCTGGTATCAGCGCGGAGACAGATGAGG 300

QY 481 C-----GAAGGAGAGAGAGAGAGAGAAAGGCTTGGGCGCTCTACAATAAATAA 533
DB 301 CGTGAAGAGGAGGAGGAGGAGGCTCTGGGCTTCTCTGCAAAAATAAATAAATAA 360

QY 534 AAAAAAAATTTAA--ATATAAATCCTATATCC-----CATATAAGATATAA 583
DB 361 TAAATAAATTTTAAAAATAAATAAATCACTATATACACATATAAAGAAATAA 420

QY 584 AGAGTCTCAGT---GCAGTATTGGCAAAATTAATCCATTTCTTTTAAATACGGG--- 635
DB 421 GAAGTCTCAGTGGCAGCTATTGTCAAAATTAATCCATTTCTTTTATATACGGTGAA 480

QY 636 AATAATGGCAATTATAGATCTGAAATTTGA-CCACTTAATGAAGCGGC-ACCCAGGTGTT 693
DB 481 TATTGGCAATTATAGATCTGGAATTTGAACCACTTAATGAAGCGGCAACACAGGTATT 540

QY 694 TTGAGGTGTTCGCATCTTCGCTGATTGGCTGTTCCCAATCTTTACATTTTATCTTT 753
DB 541 TTGAGGTGTTCGCATCTTCGCTGATTGGCTGTTCCCAATGTTTACATTTTATCTTT 600

```

ORIGIN:

```

QY 754 GCAAAAATGGTCTGTGTCACCTGGATGGAATGCTGTCCAGTTTATTTTATGTT 813
DB 601 GCAAAAATGGTCTGTGTCACCTGGATGGAATGCTGTCCAGTTTATTTTATGTT 660

QY 814 GTTATCCTTGGATGTACAAAATTCAGAAAATGATCTCTGTAGATATTCGTCTTTATTT 873
DB 661 GTTATCCTTGGATGTACAAAATTCAGAAAATGATCTCTGTAGATATTCGTCTTTATTT 720

QY 874 TGGTCATCTTTAGAAATGTTATCAGGAATGTTTAAAAACAAGAAAGAGAACTTTTCTAAGA 933
DB 721 TGGTCATCTTTAGAAATGTTATCAGGAATGTTTAAAAACAAGAAAGAGAACTTTTCTAAGA 780

QY 934 ATGATACATAGAAAAGA-----TTTTATTTTAAATGAGTTGTTAAAGCTTGTTCTTT 987
DB 781 ATGATACATAGAAAAGATCCTTTTATTTTAAATGAGTTGTTAAAGCTTGTTCTTTCT 840

QY 988 TGTTCCTGCAAGCTATCTGCCCAAGTTAATGCAATGACACATTTTATATCTCAGAAA 1047
DB 841 TGTTCCTGCAAGCTATTTGCCCAAGTTAATGCAATGACACATTTTATATCTCAGACAC 900

QY 1048 ACACACACACACACACACACACACACACACACACACACACACACACACACACAC 1100
DB 901 CCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACACACACACACAC 953

```

RESULT 6

CN646790

LOCUS

DEFINITION

5' similar to Bases 1 to 360 highly similar to human CamKIIalpha

(Hs.197922), mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE:

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

ILLUMIGEN BIOSCIENCES INC.

2203 AIRPORT WAY S, SUITE 450, SEATTLE, WA 98134, USA

TEL: 2063780400

FAX: 2063780408

EMAIL: CMAGNESS@ILLUMIGEN.COM

SEQUENCED ON 2004.03.11. 819 Q20 BASES.

PCR PRIMERS

FORWARD: CCTCACTAAAGGGAACAAA

BACKWARD: CACTATAGGCGCAATTGGGTA

INSERT LENGTH: 950 STD ERROR: 0.00

PLATE: CL000143 ROW: C COLUMN: 01

SEQ PRIMER: CCTCACTAAAGGGAACAAA

POLYA=Yes.

Location/Qualifiers

1..950

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

/clone="IBIUM:8247"

/sex="female"

/dev_stage="adult"

/lab_host="E. coli SOLR"

/clone_lib="Katze MMR"

/note="Organ: Brain; Vector: Uni-ZAP XR; Site 1: Ecor I; Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)"

ORIGIN:

Query Match	55.4%	Score 611.2	DB 7	Length 950
Best Local Similarity	87.3%	Pred. No. 1.9e-119		
Matches 829	Conservative 0	Mismatches 83	Indels 38	Gaps 13
QY	185	CAACTTCTTCGGCGCGCGGAGAACAAACGGCGCGCCCAAGCTGGGCCAGATCGGCCGAG	244	
DB	1	CAACTTCTTCGGCGCGCGGAGAACAAACGGCGCGCCCAAGCTGGGCCAGATCGGCCGAG	60	
QY	245	CAAGCGGCTGTATTATGAAGATGATGAGGATTTGATGACGTCTGTAATAATATGACCGACA	304	
DB	61	CAAGCGGCTGTATTATGAAGATGATGAGGATTTGATGACGTCTGTAATAATATGACCGACA	120	
QY	305	GGCACT-CTGGTGTCTAACT-CCCAAGAACAATGAGTCTTAAGGAGAGAAATAGGAACGG	362	
DB	121	GGCACTCTGGTGTCTAACTCCCAAGAACAATGAGTCTTAAGGAGAGAAATAGGAACGG	180	
QY	363	CGGTAACAGTTATTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTTACTAGC	422	
DB	181	CGATAACAGTTATTGGCAAAAGCATGAAAGAGAAAGCACTTTGAAATTTTACTAGC	240	
QY	423	TTG-TACCCACGATGAATCAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGG	480	
DB	241	TTGCTACCCACGATGAATCAACAACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGG	300	
QY	481	C-----GAAGGAGGAAGAGAGAGAAAGGCTTGGGCCCTCTACAAATAAATAA	533	
DB	301	CGTGAGAGGAGGAGGAGGAGGAGGCTCTGGGCTCTCTGCAAAAATAAATAAATAA	360	
QY	534	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	583	
DB	361	TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	420	
QY	584	AGAGTCTCAGT- ---GCAGTATTGGCAAAAATAAATCCATTTCTTTTATACGGGAATA	639	
DB	421	GAAGTCTCAGTTCGAGCTATTGTCAAAAATAAATCCATTTCTTTTATACGGTAA	480	
QY	640	TT-----GGCATTATAGATCTGGATTTTG-ACCATTAAATGAAGCGGC-ACCACAGGTGTT	693	
DB	481	TATTGGCGAATATTAGATCTGGATTTTGAACCATTAATGAAGCGGCAACACAGGTATT	540	
QY	694	TTGAGGTGTGGCACTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTTAAATCTT	753	
DB	541	TTGAGGTGTGGCACTCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTTAACTTT	600	
QY	754	GCAAAAATGGTCTGTGTCACATTGGATGTGAATGCTGCCAGTTTTATTTTTTATGTT	813	
DB	601	GCAAAAATGGTCTGTGTCACATTGGATGTGAATGCTGCCAGTTTTATTTTTTATGTT	660	
QY	814	GTTATCTTGGATGTACAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTT	873	
DB	661	GTTATCTTGGATGTACAAAAATTCAGAAAATGATCTCTGTAGATATTCTGTTTTATTT	720	
QY	874	TGGTCACTCTTAGAAGTTATCAGGAATGTGTTTAAAAACAAGACAGAACTTTTCTAAGGA	933	
DB	721	TGGTCACTCTTAGAAGTTATCAGGAATGTGTTTAAAAACAAGACAGAACTTTTCTAAGGA	780	
QY	934	ATGATA CATGAAAAAGA-----TTTTATTTTTAAAAATGAGTTGTAAGCTTGTTTTCTT	987	
DB	781	ATGATA CATGAAAAAGA-----TTTTATTTTTAAAAATGAGTTGTAAGCTTGTTTTCTT	840	
QY	988	TGTTGCTGCAAGCTATCTGCCAGTTAATGCAATGGAACATTTTTTTATGTCAGAAAA	1047	
DB	841	TGTTGCTGCAAGCTATTTGGCCCAAGTTAATGCAATGGAACATTTTTTTATGTCAGAAAC	900	
QY	1048	AC	1097	
DB	901	ACACACACACACACAC-CCACCCACGCAACACACACACACACACACACACACACACAC	949	

RESULT 7

CN646555

LOCUS

CN646555

954 bp

mRNA

linear

EST 13-MAY-2004


```

REFERENCE 1 (bases 1 to 951)
AUTHORS Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magness, C.L.
TITLE Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL Unpublished (2003)
COMMENT Contact: C. Magness
Illumigen Biondices Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.03.19. 778 Q20 bases.
PCR Primers
FORWARD: CCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGAATGGGTA
Insert Length: 951 Std Error: 0.00
Plate: CL000175 row: F column: 10
Seq primer: CCTCACTAAAGGGAACAAA
POLYA=Yes.

FEATURES
source
1..951
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBIOW:8006"
/sex="female"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_lib="Katze MMR"
/note="Organ: Brain; Vector: Uni-ZAP XR; Site 1: EcoR I; Site 2: Xho I; Created from Stragene ZAP-cDNA synthesis kit (catalog #200400) and ZAP-cDNA Gigapack III Gold Cloning Kit (Catalog #200450)"

ORIGIN
Query Match 55.1%; Score 608.2; DB 7; Length 951;
Best Local Similarity 86.3%; Pred. No. 8.3e-119;
Matches 821; Conservative 0; Mismatches 93; Indels 37; Gaps 12;

QY 185 CAACCTCTTCGCGCGGCGGAGCAAGCGCGCGCCCAAGCTGGCGCAGATCGCGCGGAG 244
DB 1 CAACCTCTTCGCGCGGCGGAGCAAGCGCGCGCCCAAGCTGGCGCAGATCGCGCGGAG 60

QY 245 CAAGCGGGTGTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATACCGACAA 304
DB 61 CAAGCGGGTGTATTGAAGATGATAGGATTGATGACGTGCTGAAAAATATACCGACAA 120

QY 305 GGCACCT-CTGTGTCTTAAC-CCCCAAGACAATGAGTTAAGGAGAGAGATAGGAACGG 362
DB 121 GGCACCTCTGTGTCTTAAC-CCCCAAGACAATGAGTTAAGGAGAGAGATAGGAACGG 180

QY 363 CGGTAAACAGTTATTGCAAAAGCATGAAAGAGAAAGCATTCTGAAATTTATTACTAGC 422
DB 181 CGATTAACAGTTATTGCAAAAGCATGAAAGAGAAAGCATTCTGAAATTTATTACTAGC 240

QY 423 TTG-TACCCACGATGAATCAACACCTGTATCTGTAT-ATGCCCGGAGACAGATTAGG 480
DB 241 TTGCTACCCACGATGAATCAACACCTGTATCTGTATCTGTATCAGCGCGGAGACAGATGAGG 300

QY 481 C-----GAAGGAGAGAGAGAGAGAGAAAGGCTTGGGCCCTCTTACAAATAAATAA 533
DB 301 CGTGAGAGAGAGAGAGAGAGAGGCTCTGGGCTTCTCTGCAAAATAAATAAATAA 360

QY 534 AAAAAAATTTAA--ATAATAATCCCTATATCC-----CATATAAGATATAA 583
DB 361 TAAATAAATTTTAAAAATAAATAAATCACTATATACACATATAAAGAAATAAATAA 420

QY 584 AGAGTCTCAGT----GCAGTATGGCAAAATTAATCCATTCTCTTTTAAATACGGGAATA 639
DB 421 GAAGTCTCAGTGGCAGCTATTGTCAAAATTAATATCCATTCTCTTTTATATACGGTGA 480

QY 640 TT-----GGCATTATAGATCTGATTTTG--ACCATTAAATGAAGCGGC--ACCCAGGTGT 693

```

```

DB 481 TATTGGCAATATATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACCGATT 540
QY 694 TTGAGGTGTGGCATTCTTCGCTGATTTGGCTGTTCCTCAATGTTTACATTTAATCTT 753
DB 541 TTGAGGTGTGGCATTCTTCGCTGATTTGGCTGTTCCTCAATGTTTACATTTAATCTT 600
QY 754 GCAAAAATGGTTCGTGTCACCTTGGATGGAATGCTGTCCAGTTTATTTTATGTT 813
DB 601 GCAAAAATGGTTCGTGTCACCTTGGATGGAATGCTGTCCAGTTTATTTTATGTT 660
QY 814 GTTATCCTTGGATGTACAAAAAATTCAGAAATGATCTCTGTAGATATTTCTGTTTATTT 873
DB 661 GTTATCCTTGGATGTACAAAAAATTCAGATATGATCTCTGTAGATATTTCTGTTTATTT 720
QY 874 TGGTCATCTTTAGAGTTTATCAGGAATGTGTTTAAAAACAAGAGAGAACTTTTCTAAGA 933
DB 721 TGGTCATCTTTAGAGTTTATCAGGAATGTGTTTAAAAACAAGAGAGAACTTTTCTAAGA 780
QY 934 ATGATACATAGAAAAAGA-----TTTTATTTTAAATGAGTTGTAAGCTTGTGTTCTT 987
DB 781 ATGATACATAGAAAAAGATCCTTGTGTTTATTTTAAATGAGTTGTAAGCTTGTGTTCTT 840
QY 988 TGTGTCTCAAGCTATCTGCCCAAGTTAATGCAATGACACATTTTATGTCAGAAAA 1047
DB 841 TGTGTCTCCAGCTATTGTCGCCAGTTAATGCAATGACACATTTTATGTCAGAAAC 900
QY 1048 ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 1098
DB 901 CCACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 951

RESULT 11
LOCUS CR604926
DEFINITION full-length cDNA clone CS0DF038YH05 of Fetal brain of Homo sapiens (human).
ACCESSION CR604926
VERSION CR604926.1 GI:50485733
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1566)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 1566)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
source
1..1566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF038YH05"
/tissue type="Fetal brain"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 55.0%; Score 607.2; DB 3; Length 1566;
Best Local Similarity 89.2%; Pred. No. 1.4e-118;

```

Matches	787;	Conservative	0;	Mismatches	63;	Indels	32;	Gaps	11;
Qy	252	GTGTGTTATTGAAGATGATAGCAATGATGACGTCTGAAAAATATGACCGACAAGGCACCT	311						
Db	279	GTGTGTTATTGAAGATGATAGCAATGATGACGTCTGAAAAATATGACCGACAAGGCACCT	338						
Qy	312	-CTGCTGTCTAACT-CCCCAAGACAATGAGTTAAGGGAGAGAAATAGACGCGCGGTAAAC	369						
Db	339	CCTGGTGTCTAACTCCCCCAAGACAATGAGTTAAGGGAGAGAAATAGACGCGCGGTAAAC	398						
Qy	370	AGTTATTGGCAAAAAAGCATGAAAAAGACAATTTGAAATTTATTACTAGCTGTG-TAC	428						
Db	399	AGTTATTGGCAAAAAAGCATGAAAAAGACAATTTGAAATTTATTACTAGCTGTGCTAC	458						
Qy	429	CCAGATGAATAATCAACAACCTGTATCTGCTAT-ATGCCCGGAGACAGATTAGGCGA-AGG	486						
Db	459	CCACGATGAATAATCAACAACCTGTATCTGCTATCTAGGCGCGGAGACAGATGAGCGCAGAGG	518						
Qy	487	AGGAAGAGAGAGAGAGAAGAGGCTTGGGC-CCTCTCAAAATATAAATAAAAAAATAAATT	545						
Db	519	AGGAGGAGGAGGAGAGAGAGGCTCTGGGCTCTCTGCATAAATAAATAAATAAATAAATAA	578						
Qy	546	TAAATTAATAAATAATCCCTATA------TCCCATATAAGAAATAAAAGAGTC	589						
Db	579	TAAATTTTAAATAATAAATAATTCACTATATACATATAAGAAATAAATAAAGAGATC	638						
Qy	590	TCAGT----GCAGTATTGGCAAAATTAATTCATTCTCTTTTATAACGG----AATATT	641						
Db	639	TCAGTTCGACGCTATTGTGCAAAATAATATCCATTCTCTTTTATATACGTTGAATATGC	698						
Qy	642	GGCAATTATAGATCTGGATTTTG-ACCACCTTAATGAAGCGGC-ACCACGAGTGTTTTGAGG	699						
Db	699	GCAATTATAGATCTGGATTTTGAACCACTTAATGAAGCGGCAACACACAGGTGTTTGAGG	758						
Qy	700	TGTTGGCAATCTTCGCTGATTTGCTCTGTTCCCAATGTTTACATTTTATCTTGCAAAA	759						
Db	759	TGTTGGCAATCTTCGCTGATTTGCTCTGTTCCCAATGTTTACATTTTATCTTGCAAAA	818						
Qy	760	ATGTTCTCTGTGCATCTGGATGTGAAATGCTGTCCAGTTTATTTTTTTTATGTTGTATC	819						
Db	819	ATGTTCTCTGTGCATCTGGATGTGAAATGCTGTCCAGTTTATTTTTTTTATGTTGTATC	878						
Qy	820	CTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTCGTGTTTATTTTGGTCA	879						
Db	879	CTTGGATGTACAAAAATTCAGAAAAATGATCTCTGTAGATATTCGTGTTTATTTTGGTCA	938						
Qy	880	TCCTTTAGAAGTTATCAGGAATGTGTTTAAAAACAAGAGAGAACTTTTCTAAGGAATGATA	939						
Db	939	TCCTTTAGAAGTTATCAGGAATGTGTTTAAAAACAAGAGAGAACTTTTCTAAGGAATGATA	998						
Qy	940	CATAGAAAAGATTTTATTTTAAATGAGTTGTAAAGCTTGCTGTTCTTTTCTGCTGCAAG	999						
Db	999	CATAGAAAAGATTTTATTTTAAATGAGTTGTAAAGCTTGCTGTTCTTTTCTGCTGCAAG	1058						
Qy	1000	CTATCTGCCCAAGTTAATGCAATGGACACATTTTTTATGTCAGAAAAACACACACAC	1059						
Db	1059	CTATCTGCCCAAGTTAATGCAATGGACATTTTTTATGTCAGAAAAACACACACAC	1118						
Qy	1060	ACACACACACACACACACACACGAAAAAATAAAAAA	1101						
Db	1119	ACACACACACACACACACACACGAAAAAATAAAAAA	1160						

RESULT 12	CN6644007	955 bp	linear	EST 13-MAY-2004
LOCUS	CN644007			
DEFINITION	ILLUMIGEN MQ0 9278 Katze MBR Macaca mulatta cDNA clone IBTUM.8814 5' similar to bases 1 to 360 highly similar to human Cankiinalpha (Hs.197922), mRNA sequence.			
ACCESSION	CN644007			
VERSION	CN644007.1	GI:47157450		
KEYWORDS	EST.			
SOURCE	Macaca mulatta (rhesus monkey)			

Db 421 GAAGTCTCAGTTCGAGCTATTGTGCAAAATTAATATCCATTCTTTTATATACGGTGA 480
Qy 636 AATATTGGCAATPATAGATCTGGATTTTGA-CCACTTAATGAAGCGC-ACCCAGGTGT 693
Db 481 TATTGCGCAATTTATAGATCTGATTTTGAACCACTTAATGAAGCGCAACACAGGTATT 540
Qy 694 TTGAGGTGTGGCATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTT 753
Db 541 TTGAGGTGTGGCATCTTCGCTGATTTGGCTGTTCCCAATGTTTACATTTAATCTT 600
Qy 754 GCAAAATATGTTCTGTGCACTTGGATGTGAATGCTGTCAGTCTTTTATTTTATGTT 813
Db 601 GCAAAATATGTTCTGTGCACTTGGATGTGAATGCTGTCAGTCTTTTATTTTATGTT 660
Qy 814 GTTATCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATCTGTTTTATT 873
Db 661 GTTATCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATCTGTTTTATT 720
Qy 874 TGGTCATCTTTAGAATTTATCAGGAATGTGTTTAAACAGAAAGAACTTTTCTAAGA 933
Db 721 TGGTCATCTTTAGAATTTATCAGGAATGTGTTTAAACAGAAAGAACTTTTCTAAGA 780
Qy 934 ATGATACATAGAAAAGA-----TTTATTTTAAATGAGTTGTAAGCTTGTGTTCTT 987
Db 781 ATGATACATAGAAAAGATCTTGTGTTTATTTTAAATGAGTTGTAAGCTTGTGTTCTT 840
Qy 988 TGTGCTGCAAGCTATCTGCCAAGTTAATGCAATGGAACATTTTTTATGTCAGAAA 1047
Db 841 TGTGCTGCAACTTATTTGCCCAAGTTAATGCAATGGAACATTTTTTATGTCAGAA 900
Qy 1048 AC 1101
Db 901 ACACCCACCCCCCCACCCACCCCCACCCACCCACACACACACACACACACACACAC 954

RESULT 13

CN803392
LOCUS
DEFINITION
ILLUMIGEN MCQ 32916 Katze MBR Macaca mulatta cDNA clone
IBIUW:14140 57 similar to Bases 61 to 924 highly similar to human
CamKIIalpha (Hs.197922), mRNA sequence.

ACCESSION

CN803392.1 GI:47699368

VERSION

EST.

KEYWORDS

Macaca mulatta (rhesus monkey)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;

Cercopitheciinae; Macaca.

1 (bases 1 to 954)

Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.

Large-scale Rhesus Macaque cDNA Sequencing

Unpublished (2003)

Contact: C. Magness

ILLUMIGEN Biosciences Inc.

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA

Tel: 2063780400

Fax: 2063780408

Email: cmagness@illumigen.com

Sequenced on 2004.04.03. 803 Q20 bases.

PCR Primers

FORWARD: CCCTCACTAAAGGGGAACAAA

BACKWARD: CACTATAGGCGAATTGGTA

Inert Length: 954 Std Error: 0.00

Plate: CL000275 row: C column: 01

Seq primer: CCCTCACTAAAGGGGAACAAA

POLYA=Yes.

FEATURES

source

Location/Qualifiers
1..954

/organism="Macaca mulatta"

/mol_type="mRNA"

/strains="Indian"

/db_xref="taxon:9544"

/clone="IBIUW:14140"

/sex="female"

/dev_stages="adult"

/lab_host="E. coli SOLR"

/clone_lib="Katze MBR"

/note="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I;

Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis

kit (catalog #200400) and ZAP-CDNA Gigapack III Gold

Cloning Kit (Catalog #200450)"

ORIGIN

Query Match 54.8%; Score 605.4; DB 7; Length 954;
Best Local Similarity 86.0%; Pred. No. 3.3e-116;
Matches 820; Conservative 0; Mismatches 96; Indels 37; Gaps 12;
Qy 185 CAATCTTCTTCGGCGCGGCGAGAACAGCGCGCGCCCAAGCTGGGCCAGATCGGCCGAG 244
Db 1 CAATCTTCTTCGGCGCGGCGACACACAGCGCGCGCCCAAGCTGGGCCAGATCGGCCGAG 60
Qy 245 CAAGCGGTTGTTTATGAAGATGATAGGATGTAGCTGTGAAATATATGACCGACAA 304
Db 61 CAAGCGGTTGTTTATGAAGATGATAGGATGTAGCTGTGAAATATATGACCGACAA 120
Qy 305 GGCACCT-CTGGTGTCTAACT-CCCCAAGACAATGAGTTAAGGAGAGAATAAGAACGG 362
Db 121 GGCACCTCTCTGGTGTCTAACTCCCCAAGACAATGAGTTAAGGAGAGAATAAGAACGG 180
Qy 363 CGGTAACAGTTATTGGGCAAAAAGCATGAAAAGAGAAAGCACCTTTGAAAATTTATTACTAGC 422
Db 181 CGATAACAGTTATTGGGCAAAAAGCATGAAAAGAGAAAGCACCTTTGAAAATTTATTACTAGC 240
Qy 423 TTG-TACCCACGATGAATCAACACCTGTATCTGGTAT-ATGCCCGGAGACAGATTAGG 480
Db 241 TTGTTACCCACGATGAATCAACACCTGTATCTGGTATCAGCGCGGAGACAGATTAGG 300
Qy 481 C-----GAAGGAGGAGAGAGAGAGAAAGCTTTGGCCCTCTACAAATAAATAA 533
Db 301 CGTGAGAGAGGAGGAGGAGAGGCTCTGGCTTCTCTGCAAAAATAAATAAATAAATAA 360
Qy 534 AAAAAAATAATTTAAA---ATAATAAATCCTATATCC-----CATATAAGATAAATA 583
Db 361 TAATAAATAATTTAATAATAATAAATAATCACTATATACACATATTAAGAAATAAATA 420
Qy 584 AGAGTCTCAGT----GCAGTATTGCGCAAAATTAATCCATTCTTTTTTAATACGGG- 635
Db 421 GAAGTCTCAGTTGCAGCTATTGTCAAAATTAATATCCATTCTTTTATATATACGGTGA 480
Qy 636 AATATTGCAATATAGATCTGATTTTGA-CCACTTAATGAAGCGGC-ACCCAGGTGT 693
Db 481 TATTGCGCAATTTATAGATCTGATTTTGAACCACTTAATGAAGCGGCAACACAGGTATT 540
Qy 694 TTGAGGTGTTGGCATTTCTCGCTGATTTGGCTGTTCCCAATGTTTACATTTATTAACTTT 753
Db 541 TTGAGGTGTTGGCATTTCTCGCTGATTTGGCTGTTCCCAATGTTTACATTTATTAACTTT 600
Qy 754 GCAAAAAATGTTTCTGTGACATTTGGATGGAATGCTGTCCAGTTTATTTTATTTATGTT 813
Db 601 GCAAAAAATGTTTCTGTGACATTTGGATGGAATGCTGTCCAGTTTATTTTATTTTATGTT 660
Qy 814 GTTATCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATCTGTTTTTATTT 873
Db 661 GTTATCTTGGATGTACAAAAATTCAGAAATGATCTCTGTAGATATCTGTTTTTATTT 720
Qy 874 TGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAACAGAAAGAACTTTTCTAAGGA 933
Db 721 TGGTCATCTTTAGAAGTTATCAGGAATGTGTTTAAACAGAAAGAACTTTTCTAAGGA 780
Qy 934 ATGATACATAGAAAAGA-----TTTTATTTAAATGAGTTGTAAGCTTGTGTTCTT 987
Db 781 ATGATACATAGAAAAGATCTTGTGTTTATTTTAAATGAGTTGTAAGCTTGTGTTCTT 840
Qy 988 TGTGCTGCAAGCTATCTGCCAAGTTAATGCAATGGAACATTTTTTATGTCAGAAA 1047

Search completed: March 30, 2005, 03:06:42
Job time : 3521 secs

188	Qy	CTTCTTCGGCGCCGGCGCAGAACAGCGCGCCGCCAGCTGGGCCNGATGCGGCCGGAGCAA	247
13	Db	CTTCTTCAGCGCCCGCCCTCAAACGAGCGCGCGCCCAAGCTGTGTCCAATCTCGCGCGGAGCAA	72
248	Qy	GCGGGTTGTTATTGAAGATGATAGGATTTGATGACGTGCTGAAAAAATATGACCGACDAGGC	307
73	Db	GCGGGTTGTTATTGAGAGTGTATAGGATTTGATGACGTGCTGAAAAAATATGACCGACDAGGC	132
308	Qy	ACCT-CTGTGTCTTAACCT-CCCCAAAGACAATAGAGTTAAGGGAGAGAAATAGGAAACGGCGG	365
133	Db	ACCTCTCTGTGCTTAACCTCCCCAAAGACAATAGAGTTAAGGGAGAGAAATAGGAAACGGCGA	197
366	Qy	TAAACAGTTATTGGCAAAAAGCATGAAAAGAGAGAAAGCACTTTGAAATTTTATTACTAGCTTG	425
193	Db	TAAACAGTTACTGGCAAAAAGCATGAAAAACGAAAAGCACTTTGAAATTTTATTACTAGCTTG	252
426	Qy	-TACCCACGATGAAATCAACAACCTGTATCTGTGTAT-ATGCCCGGAGACAGATTAGGC--	481
253	Db	CTACCCACGATGAAATCAACAACCTGTATCTGTGTATCAGGCCGGGAGACAGATGNGGGT	312
482	Qy	-----GAAGGAGGAGAGAGAGAGAAAGGCTTGGGCCCTCTCAAAATAAAAATAAAAA	536
313	Db	GAGAAAGGAGGAGGAGGAGAGGCTCTGGGCTCTCTGCAAAAAATAAAAAATAAAAAATAA	372
537	Qy	AAAAAAATTTAAA--ATAATAAAATCCCTATATCC-----CATATAAGAAATAAAGA	586
373	Db	ATAAAAAATTTAAAAAATATAAAAAATCACTATATACACATATAAAGAAATAAAAAAGA	432
587	Qy	GTCTCAGT----GCAGTATTGGCAAAAATAAATCCATTTCTTTTAAATACGGG----AAT	638
433	Db	GTCTCAGTTGCAGCTATTGTGTAATAATCAATTTCTTTTATATACGGTGAATAT	492
639	Qy	ATTGGCATTTATAGATCTGGATTTTG-ACCACTTAATGAAGCGGC-ACCCACAGGTGTTTG	696
493	Db	TGCCCAATTTATAGATCTGGAATTTTGAACCACTTAATGAAGCGGCAACACCGAGTATTTG	552
697	Qy	AGGTGTTGGCATCTCTCGCTCATTTGGCTGTTCCCAATGTTTTACATTAATTTAACTCTTGA	756
553	Db	AGGTGTTGGCAATCTTGCTCATTTGGCTGTTCCCAATGTTTTACATTAATTTAACTCTTGA	612
757	Qy	AAAAATGGTCTGTGCGCATCTGGATGTGAAATGCTGTCCAGTTTTTATTTTTTTTATGTTGT	816

This Page Blank (uspto)